THESIS ABSTRACT:

INERT INSTRUCTION: AN EVALUATION OF OPPOSING LABEL, OBJECT, AND GOAL-ORIENTATION DISPLAY FORMATS IN COMMUNICATING WITH MUSEUM VISITORS

This study set out to assess the impact of selected exhibit design formats on the reactions of visitors to the National Museum of Antiquities of Scotland. The principal aim of this study was to evaluate the effectiveness of opposing object, label, and goal presentations to communicate with, that is, to attract, interest, and instruct, the casual museum visitor. The contrasting object formats were: an unstructured display versus a thematically grouped one. The text variations were: single block labels versus a series of multiple labels with bold headings. Goal statements were either present or absent from the other format combinations.

The treatment combinations were successively varied within one museum display case. They were evaluated using observation of visitor movement patterns and timed viewing, and written test scores and visitor comment responses. The results of these measurements were analysed using a comparison of treatment means, multiple regression, chi square, and correlation coefficients. The results provide conclusions regarding the use of bold headings, orientating statements, and thematic displays. Further implications are provided for cognitive assessments in museums and for visitor behavioural patterns, such as viewing sequences and use of signs.

TABLE OF CONTENTS

APPENDICES
TABLE OF CONTENTS

NOTE- the initials "M.D.T." refer to the Main Display Treatments, i.e. MT,MTG, BT,BTG, BU,BUG

LIST OF TABLES

INTRODUCTION

TO

INERT INSTRUCTION

FOR A

MOBILE AUDIENCE

The National Museum of Antiquities of Scotland has been host to a novel experiment in learning within a museum context. The aim of this experiment has been to evaluate the relative effectiveness of opposing label, object, and goal-orientation display formats in communicating with museum visitors.

A single museum display case was the vehicle for this experiment. Within the display case, selected label, object, and goal treatments were successively varied over a four month period, from April to July. Only the design formats of the labels, objects, and goals were changed in each treatment. The display case contents, both objects and label wording, and its background, remained constant throughout the main experiments.

The opposing label layouts were: 'blocks' of main label text similar in format to the pages of a book, versus 'multi-track' text characterized by newspaper-like headings and spatially separated paragraphs. The objects were either placed in an 'unstructured' layout, or grouped into one of four, horizontal, 'thematic' bands. Goals, when they were present, were displayed as orientating statements on the sign directing visitors to the display case, and as thought provoking questions at the beginning and end of the case. Within each of the successive display treatment variations, one of the label formats would be shown with one of the object layouts, and they would appear with or without the presence of goal statements.

Four measurements were utilized to evaluate the communication effectiveness of the opposing object, label, and goal formats. The first two measurements were taken from the observation and timing of the reactions of all of the visitors who walked into the gallery. These global measures were: the display treatment's ability to attract the visitor population, and then, to retain their interest. The two other measurements were taken from the written responses of visitor samples who were asked to complete a questionnaire after they had casually looked at the display case. These measurements evaluated the samples' knowledge gain and their written comment responses.

To present the focus of this study in operational terms, the following questions were posed:

- **Labels**– Can short text and the use of bold headings attract visitors, or increase their attention spans, or their ability to assimilate information?
- **Objects**– If the material in a display case is organized appropriately, will many visitors attempt to follow themes from one section of the display case to another?
- **Goals**– Do orientating statements (questions, goals, aims) increase visitor viewing times, knowledge gains, or the number of favourable comments?

This study was instigated in order to provide information of relevance to the museum profession. It addresses their frequently voiced concern at having insufficient information about the relative effectiveness of static, low-budget display techniques in communicating with museum visitors. Answers to the questions above should help museum professionals to choose display techniques which increase their effectiveness in communicating concepts, facts, and 'wonderment' to their audience. From the outset it was hoped that the knowledge gained from this study would prove useful in formulating guidelines for displays both within this museum and within others which have a similar type of collection.

The study presented here is unique in its selection of the experimental variations tested and in the detail of cross-linked information obtained from the results of a test as well as from detailed observational and demographic data. Due to the dearth of relevant studies, many fundamental questions needed to be answered. Although not the main focus of the study, such answers should provide the museum staff with additional important clues concerning visitor actions, reactions, and characteristics. Several of the generally important premises under scrutiny were:

- Is it possible that the person just off the street will be able to learn anything at all while standing up in a potentially noisy, crowded, unfamiliar environment?
- What length of time can visitors realistically be expected to view a display case?
- What is the impact of viewing sequence on maintaining interest and facilitating comprehension?
- Will visitors utilize directional devices such as signs?

In order to provide an overview of the content and layout of this thesis, a summary of each chapter is given below.

Chapter 1 begins with a review of the relevant literature. A synopsis is provided of previous museum visitor research. Special emphasis is given to other museum studies which evaluated contrasting display techniques. Attention is also given to a discussion of labelling practices.

Chapter 2 presents a general overview of the design of this research. Following a brief summary of the research design practices used, the context of the study is explained. This first section covers the initial contact with the museum staff, a description of the museum and its collection, the selection of a display case as the vehicle for the study, the time frame and population sampled, and the restrictions on the choice and arrangement of objects and

labels. The next section provides an overview of the experimental treatment variations, including the control groups and concurrent experiments, and a brief explanation of the measurements used to evaluate the communication effectiveness of the treatment variations.

Chapter 3 deals with the organization of the display treatment variations. This chapter discusses in detail the reasons behind the selection of the particular display case used in this experiment, its format, and the selection and description of the particular object, label, and goal variations tested within it.

Chapter 4 describes the observation and testing measurements which were employed to evaluate the display treatments. Both the techniques and the tools which were used to measure the visitor responses are discussed in detail.

Chapter 5 describes the environmental conditions under which the questionnaire was administered and the demographic questions which were asked, and presents the relevant results. A general chapter summary is followed by a discussion of the implications of these results for all of the visitors to this museum, as well as a comparison with the results from other museum studies. The chapter concludes with a discussion of the probable impact on the evaluation measurement findings, of the differences between the test conditions and demographic responses for each of the display treatments.

Chapter 6 discusses the results of each of the four communication evaluation measurements in turn. The results of both the main display treatments and the control groups are dealt with and the implications of these results are stated.

Chapter 7 begins with a brief summary of the research methods, aims, hypotheses, and significant results of the study. The major conclusions and

their implications for museum display design, test and evaluation methods, and future studies, are then discussed.

CHAPTER 1

REVIEW OF THE

RELEVANT LITERATURE

1.1. PREVIOUS MUSEUM VISITOR RESEARCH

Museum visitors in Britain have primarily been recognizable only as numbers in a turnstile register. Although British museums generally espouse lofty educational aims, they have rarely attempted to identify the recipients of their wisdom or to evaluate their established teaching methods and learning environments. Therefore, museums are remarkable among publicly funded, major British institutions with primary educational aims, in the lack of research done on them. In the past, very few museums in Britain have undertaken a survey of any kind and only recently have a few attempted any measure of visitor learning.

In the United States, however, the survey tradition in museums was firmly established by the 1930's. Many of the early surveys were essentially observational studies focused on visitor actions such as "traffic" movement patterns and the time spent viewing exhibits and labels. Pioneer observers such as Robinson (1928) found that large type, short paragraphs, and nontechnical vocabulary were of equal importance when measuring the average viewing time of a display case. His contemporary, Melton (1933), found that visitors in American museums displayed a marked tendency to turn to the right when entering a room and tended to move faster as they got proportionally closer to the exit in a room. Porter's Peabody Museum Survey (1938) suggested that visitors there spent spent on average 3 minutes and 11 seconds looking at an exhibit and during that time read four labels.

In recognition of the educational need to identify the precise audience one

wishes to instruct, as well as to ascertain some of their motives, the second generation of museum surveys concentrated on obtaining descriptive data about visitor characteristics and preferences by directly approaching the visitors to a particular museum. Visitors were then interviewed or, more usually, asked to complete a questionnaire. The visitors were typically asked to supply information about their age, sex, place of residence, educational level attained, occupation, and areas of exhibit interest. Major demographic surveys were carried out by Abbey and Cameron (1959-61) in the Royal Ontario Museum in Toronto, by Philip Doughty in the Ulster Museum (1968), and by Dr. Arthur Niehoff (1953 and 1968) in the Milwaukee Public Museum, among others.

In an interview situation, attempts were also made to establish some measure of the visitors' background understanding and knowledge of the fields covered by the museums' collections. Such attempts to establish basic pre-visit knowledge were considered particularly important by those in the museum field such as de Borhegyi (1963) who advocated the use of varying levels of complexity in a museum exhibit. Others, such as Lee Parsons (1965), soon regretted not establishing such base levels when it was found that his comprehension test for visitors who had seen a new display case could be answered equally well by visitors who had never seen the test display.

Much of the recent interest in America, however, has been focused on the museums' obligations to its public. One of these obligations of particular importance is its educational role in providing an effective informal learning environment. Assessment techniques have shifted as well to take account of recent trends in education, sociology, and anthropology. More emphasis is now given to visitors' opinions. According to Alma Wittlin in her book, The Museum: Its History and Function in Education, the new educational emphasis in museums world-wide stems from the American influence in the democratization of museums away from the nineteenth century European

concept of museums for the "educated-classes" alone.

Within the last twenty years, an increasing amount of both formative, or ongoing, and summative evaluation studies are being attempted both in Britain and America (Alt, 1978; de Borhegyi, 1968). In the Smithsonian the results of 'naturalistic inquiry' research, based on interviews with visitors and observations of their movement patterns (Wolf and Tymitz, 1978), have been used to effect changes in exhibits. Additionally, attempts have also been made to determine whether viewing an exhibit could change some visitors' attitudes to strongly held beliefs (Shettel, 1966).

1.2. MUSEUM EXHIBIT EVALUATION TESTS

> "...a tightly controlled statistical analysis of visitor response to a series of variations...re concepts vs. specimens, extensive vs. minimal labeling..(etc)..and similar studies...should be the background of future exhibit planning and programming. Without them museums could easily fail in their function of mass education and become either glorified warehouses, recreational facilities, or exclusive clubs for the learned." (de Borhegyi, 1968, p.7)

The use of experimental techniques to examine the effects on museum visitors of changing specific components of a display is a relatively "untrampled" approach. Most controlled experiments involving museum visitors have been approached through the use of mock-up situations in a laboratory (Shettel, 1968) and the observation of visitors in different settings (Melton, 1935). Few people however have experimented with the physical manipulation of exhibit variables in a museum display case. Because of the scarcity of such important studies, two of them are explored in depth within this section.

Parsons (1965), used a single museum exhibit display case in the Milwaukee Public Museum to test the use of a 'question approach' in labelling, 'visual complexity' versus simplicity, and 'colour variations' versus black and white.

However Parsons' test, his principal evaluative tool, proved inconclusive and he was forced to use rating statements from visitors to assess the relative merits of his exhibit variables. This highlights the major difficulties encountered in most attempts to assess visitor learning in museums: the testing procedures suffer from a lack of adequate controls, such as piloting, demonstrable validity, and pre-selection criteria for visitor samples, and the tests themselves suffer from the problem of trying to measure "unique" learning after a brief exposure time.

Parsons used a display case 10 feet long, 7 feet high and 1 foot deep to display his design variations. 2,000 visitors filled out his questionnaire and three out of every five visitors who were asked to participate did so. However, despite all the obvious effort which went into the design of his display variations, his "quiz" and his experimental procedures have major flaws. Visitors who had not seen his test display case actually scored, on average, slightly higher than those who had seen his display and then taken the same quiz, which was designed to test the amount of knowledge gained from the test display! Thus the quiz was rendered invalid as a measure of knowledge gain since it tested knowledge which visitors either already knew or could easily guess. He kept no record of the number of visitors who did not look at his test case, so there can be no indication if a particular variation discouraged all but the most enthusiastic visitors. One-third of his sample were asked to look at the test display and consequently knew they would be tested (this was done to collect the requisite sample number in a relatively short length of time). Also, he only timed half of the visitors in the sample for his first variation, so no comparisons can be drawn.

His use of subjective exhibit "success" ratings was then the only reasonably reliable evaluative instrument left to him. He asked his visitors to mark their most appropriate response (out of three choices) to various factors of his test

exhibit case. They were asked their opinion on the exhibit topic, i.e. was it presented in an understandable manner; did they learn anything new from the exhibit; how much did they enjoy the exhibit; were the labels too detailed; was it too colourful; were too many utensils shown; were too many Indian areas shown; and "other criticism...".

Parsons concluded that museum personnel pitch their exhibits at too low a level, since their audiences are obviously better educated about the subjects on display than was anticipated. He also felt that the rating score for his 'question' approach in displays was approximately equal to that of the more usual 'declarative' type of labels. Consequently he suggested that the 'question' approach could be used occasionally to add variety and stimulate individual thought. He noted a strong preference for relative 'visual complexity' as opposed to the then fashionable 'visual simplicity', displaying few objects. He furthermore felt that too much colour could compete with the objects for attention and that whereas most visitors marked down the 'black and white' for being too plain, it scored well on every other factor and had the possible benefit of giving more emphasis to the objects.

Following on from Parsons, also undertaking a controlled testing experiment by varying single factors within a specially designed exhibit case, was Thomas Abler. He also conducted his experiment in the Milwaukee Public Museum and published his results in 1968. Abler tested his hypothesis that the direction from which a visitor approached a case, or the order in which the exhibit materials were viewed, would affect the amount learned from a museum exhibit case.

The subject matter in his test case was a sequence of Indian bag weaving, which was displayed in four stages from start to finish. His aim was to test whether there would be any significant difference in learning if the sequence

was viewed from start to finish, as opposed to being viewed from finish to start. He also aimed to discover whether or not any learning differences occurred, if visitors viewed the test exhibit case from left to right, as opposed to viewing it from right to left. Abler tested every combination of these four factors, i.e. approach from the right and view the sequence in its proper order, approach from the left and view the proper order, approach from the left and see the reverse order, approach from the right and see the reverse order. However, there was a major difficulty, which he did not control for, or assess, in his experimental design. The visitors who approached from the left saw the test display as they entered the gallery, whereas, those who approached from the right saw the case as they were leaving the gallery, and consequently were possibly tired of reading labels. This prior movement difference means that the different samples saw the test displays under differing conditions, which were not controlled for in the experimental design.

Abler used two test measures: timed viewing of the visitors in front the display case, and a short questionnaire. He recorded the times for all visitors who looked at the case for 10 seconds or longer. However he failed to note both the number of visitors present who did not look at the case, as well as those who viewed the test display case for less than 10 seconds. The lack of such measures throws doubt onto the validity of his study. It could be that the majority of visitors present refused to spend time looking at a display in, for example, reverse order, leaving only the exceptionally motivated visitors to view it, with the consequent likelihood that they would score highly on any subsequent test.

The questionnaire was divided into five parts. The first, at the top of the page, contained spaces for the tester to fill in to indicate the date, the display case variation being tested, and the time the visitor spent in front of the case. The second section asked the visitors to supply information about their sex;

age group; the number of people with them; whether or not they discussed the exhibit case with other people as they viewed it; and three questions asking if they had known they would be tested, if they had been tested before, and if they had seen the display case before.

The third part of Abler's questionnaire contained ten multiple choice questions. The questions were designed to be difficult, testing knowledge most visitors would only know from having studied the test display case. However Abler does not state on what grounds he bases his assumption that the questions do not test common knowledge. An acceptable solution to the need to assess prior knowledge, would have been to use a control group of visitors to the museum, who had not seen the display case, and who were asked to fill in the questionnaire. If they could not answer a significant number of the questions correctly, then it would be fair to assume that the questions did not test common knowledge, or provide an easy format for correct guesswork.

The fourth portion of the questionnaire asked the visitors to put in order eight steps in making a woven Indian bag. This was, in Abler's own estimation, a very difficult question, which no visitor tested was able to answer correctly. An acceptable solution to this difficulty might have been to have interviewed a preliminary sample of people, who had looked at the exhibit and completed the questionnaire, in order to assess their actual knowledge gain and to uncover any difficulties they faced in understanding a test question or relating it to the material they had seen. Another possible solution would have been to ask a number of visitors to study the display case, knowing they would be tested on the information, in order to assess maximum learning potential. Either solution would have allowed overly difficult or misleading question to be modified or discarded.

He also noted that this question was often not completed (by as much as forty-one percent in one variation) and suggested that it was frequently left blank because it required more thought than the other questions and because it was the the last question, so visitors just tended to give up. It is possible that one or two layout changes could have been used to achieve a higher response rate on this question. It might have fared better if he had placed it in the middle of his question group, or if he had organized his questionnaire differently, so that his demographic question section came after his quiz questions. Either arrangement could have been explored through pilot studies, as could a restatement of the question, to encourage a higher response rate.

Abler concluded his questionnaire with a request for comments and criticisms but noted his disappointment that the information supplied provided no useful insights. It could be that, by providing a two inch long "blank" at the end of his question, as opposed to leaving free space below the question, many people could have inferred that their responses were meant to be confined within that space, and consequently have supplied little beyond "very interesting" type of remarks. An additional approach he might have used would have been to focus the visitor's attention by asking for comments on the 'design' of the case rather than just comments on 'the case'.

Abler's sample was taken from casual weekend visitors to the museum who were over the age of thirteen, and who looked at his display case voluntarily, that is, with out being asked to do so, for ten seconds or longer. He tested all of the visitors who fulfilled these requirements but still found that to get samples of between thirty to forty testees, he had to test each variation for approximately one month.

His justification for testing only weekend visitors was that there were not enough visitors present during the week to make testing worthwhile. However

he neglected to go into the previous survey data compiled for the museum to suggest if the weekend and weekday visitors were likely to be similar or different and to suggest whether his sample was likely to be representative, in terms of sex and age group, of the known weekend population.

He does justify his "saturated sampling" technique by stating that "...a random sample is an impossible requirement in a museum testing situation" due to the dearth of visitors present who are likely to view a particular display case. He goes on to quote Parsons' experience, in a busy area of the museum, where he had to get the attendants to direct approximately one-third of his subjects to look at his case in order to get a sample of twenty-five visitors in an afternoon. Again these studies highlight a major difficulty in testing the responses of casual museum visitors to a particular display case: the length of time needed to collect an adequate sample.

Abler concludes that people who approach an exhibit from the left learn more than those who approach from the right, regardless of whether or not the display case is organized in a forward or a backward sequence. However, as Abler himself admits, this should only be accepted as a suggestion, since the validity of his test procedures is shaky.

It is also interesting to note that the visitors' viewing times for all of his tested display variations ranged from 10 to "90+" seconds. The means for the four individual variations ranged from a low of 28.46 seconds to a high of 43.18 seconds. However, there was much less difference in the medians for the variations. These ranged from 20 to 25 seconds. Abler also concluded that there was no significant increase in learning dependent on the length of time visitors looked at the display case.

Both of the studies conducted by Parsons and by Abler are instructive as pioneering efforts to perform credible controlled experiments within a real

museum situation. Both have strong points, Parsons' well thought out display case design and his rating section, and Abler's quiz section. However both suffer primarily from a lack of sound experimental methodology. Both experiments could have been improved through the use of pilot and control studies to detect poor questions and to determine viable sample groups, as well as to establish measures of internal and external validity.

1.3. A BRIEF INDICATION OF THE DISCIPLINES UNDERLYING EXHIBIT DISPLAYS

> "the museum has come to be an educational institution and increasing energy has been devoted to a refinement of definition of its educational role, and to a search for effective technique in fulfilling this function." (Marion Mochon, 1968)

Theories from Education, Psychology, Sociology, and even Antropology, have contributed, not only to the field of visitor behaviour and exhibit evaluation, but have also made a fundamental impact on exhibit design. Since World War II, design policies in museum displays have often mimicked the public sector. Goebels' effective use of mass psychological motivation through posters and stage-managed rallies for Hitler amply demonstrated the potential for psychological manipulation with advertising propaganda. Since then, designers have been increasingly called upon to utilize the results of psychological studies in order to get across the relevant messages. Studies in sociology and education, as well as psychology have also been looked to by museum designers for concept alternatives. One of the areas of museum exhibit design which has been particularly influenced by these disciplines is the format, content, and layout of labels.

1.4. LABELS:

1.4.1. THE GENERAL ROLE OF LABELS IN MUSEUMS

The primary function of labels in a museum is to communicate with the museum visitor. Labels are used to convey ideas and information which objects alone do not or can not convey. They also provide a static "permanent" reference point within the museum which the visitor may refer to as and when s/he pleases. Many of the ideas for good label practice in this study have been influenced by the works of: North, 1957; Williams, 1960; Weiner, 1963; and the Communications Design Team at the Royal Ontario Museum, 1976.

1.4.2. ALTERNATIVES TO LABELS

Alternatives or supplements to labels may be used by certain museums or used within a particular area or context in a museum. These may be categorized into three, sometimes overlapping, groups. These are electrical/mechanical, portable, and human alternatives. Their suitability is dependent on the type and resources of different museums.

Stationary electrical or mechanical devices may be used to orient visitors to particular concepts and contexts, to demonstrate how something works, and to pose questions and indicate correct responses. Films or slides with commentary are often found in orientation areas to provide a general background to a collection, or in supplementary information areas to provide in-depth information on particular topics. Push button demonstrations and working models all provide in-depth visual information which can usually be controlled, to some extent, by the visitors. Tapes can provide audial commentary for a demonstration, provide general information, or be used to suggest an appropriate atmosphere, such as winds howling through an Ice Age exhibition. Electronic question-answer boards which light up or buzz when the

correct answer is indicated, are often very popular. Other "correct response" devices, both portable and stationary, which visitors follow using a programmed learning format, have also been used in museums for explicitly didactic purposes.

There are also portable alternatives to labels. Catalogues, guide books, and flip cards may be purchased in a museum and then taken home as a reference to a collection. Audio guides used in the museum, and normally queued to particular exhibits, provide flexible information which visitors can usually control for repetition and timing between dialogues. Information sheets, with or without hand held clip boards or paddles are particularly used by historic house and folk museums where labels would detract from the authentic atmosphere. Quiz sheets may be used by any type of museum to focus attention on particularly important areas in a collection.

Human alternatives are possibly however the most familiar and potentially the most flexible. They may provide lectures, tours, or demonstrations, and often request visitor involvement through questions.

1.4.3. LABEL TYPES

Label types are generally classified by their content and aims and by their importance in an organizational structure. Label content leads to such designations as: General, Specimen, Declarative, Provocative, Interrogative, Orientation, and Objectives or Goals. And label organization to such terms as primary, secondary, and tertiary. Label aims, apart from reflecting content, may be explanatory or affective.

1.4.4. AIMS

Labels perform a variety of functions, and consequently may be used in different ways. They may place an object in context, explain its use or implications for its society, or compare it with another item. They may be

used simply to identify an object or to reinforce, clarify, or convey additional information. Labels can explain concepts demonstrated by objects or illustrations. They can also inform, intellectually stimulate, amuse, excite, arouse curiosity, amaze, provoke questions, etcetera. in their own right. They may state explicit goals or objectives for an exhibit or indicate its overall structure and organization. Labels may also bring to light the results of recent research and focus attention on 'unknowns' and areas for additional research.

1.4.5. CONTENT OF LABELS

The following criteria were particularly influential in governing the selection and presentation of written information which appeared in the display treatments discussed in subsequent chapters. Many of these prescriptions came from the Royal Ontario Museum's handbook, Communicating with the Museum Visitor.

1. GENERAL- Labels should do more than just identify an object. They can be intrinsically interesting by providing information about who made the object, when, where, how, and why. Labels can, among other things, relate an object to other items, periods, or cultures; stimulate thought and curiosity; and impart the results of recent research.

2. STRATIFIED LABELS- Labels can be stratified through colour, size, or location, so that the content can be geared to different interest and knowledge levels, ages, and backgrounds.

3. LANGUAGE- Labels should use a simple vocabulary and style which the visitor can easily understand, often they are too erudite.

4. BALANCED VIEW- between over simplification and including all uncertainties.

5. IRRELEVANT INFORMATION- Labels should add to what the eye can see. Labels should not restate the obvious in a specific section. For example, "brown earthenware pot". Also, labels should not include distracting, irrelevant information, such as the registration number or the donor's name.

6. HUMAN INTEREST- can be aroused through use of written excerpts from the past, such as, songs, tales, poems, letters, riddles, and superstitions as well as through objects.

7. OBJECTIVES- Labels should attempt to convey overall museum aims as well as specific ones for the display case or exhibition. Such aims and objectives should be explicit rather than implicit.

8. QUESTIONS AND GOALS- should be meaningful and challenging. They should not simply ask the visitors to locate single items and facts in the exhibit. Instead, they should require the application of facts learned from the exhibit, concepts, and principles, to new situations. These types of questions and goals increase learning and retention. Good question asking labels reduce learner fatigue.

1.4.6. LABEL LOCATION AND ORGANIZATION

1.4.6.1. Diverse Opinions on Label Location

The location of labels within a display case is dependent on the labelling system adopted. There are two major labelling systems which are currently popular. In one all the labels are grouped together in the case in a separate section from the objects in the display. In the other, more conventional system, the labels are placed in relevant positions within the display, for example, next to the objects they identify.

The "grouped" system is dependent on the use of identity 'keys' which may be numbers, letters, colours, etcetera. In this system, each object is issued with a unique identifying unit, such as a number, usually located beside it, and the same number appears beside the relevant label, which is located away from the object (for example at the bottom of the case). Ideally, corresponding numbers should easily be found by the viewer. This is made easier if the location of a label bears some relationship to the location of its object within the exhibit, or to a drawing of all the objects in the display. Proponents of this system, usually designers who like everything neat and tidy, maintain that specimen labels close to the artifacts are visually distracting.

In the "adjacent" system, specimen labels are placed directly beside the objects which they describe. Distraction from the objects may be minimized by using backgrounds to the label print which are perspex or which match the background of the display. Proponents of this system maintain that visitors are less confused, and more likely to read a label, if it's beside its relevant object than if it's in a separate location. They also allege that less time is wasted searching for information. This system was used for all test variations in this study.

1.4.6.2. Additional Principles of Label Location and Organization

Following good exhibit display practice, label organization and location should reflect certain principles. Exhibit objectives and organizers should be restated within the exhibit. General questions are best located just after the material they cover. Declarative labels should be interspersed with other labels. Important labels should be located within the optimum field of vision, which is from eye level down to a few feet below it. Mean adult eye level is around five feet two and one-eighth inches (5'2 1/8"). Labels should be tilted up if they are well below that level.

1.4.6.3. Label Organization for Handicapped Visitors

Where possible and appropriate, special provision should be made for handicapped visitors. Those in wheel-chairs are helped if upper shelves are transparent and provide duplicate labels which are tilted so that they can be read from below. For those who are blind, Braille could be used where appropriate for short labels, not much longer than thirty words.

1.4.7. LABEL FORMATS

Studies of text reading ability and the observation of museum visitors' label reading habits suggest a number of label format principles. Concise labels are almost universally recommended. Results from observation studies anticipate a

short viewing time. They suggest an upper limit of seventy-five (75) words in the main text. Instead of words, illustrations, charts, diagrams, cutaway models, etcetera. could be used. Text should be divided into several paragraphs instead of one long one. Long passages should be divided into several sentences and the line lengths of sentences should be kept short. Words should also be short and easy to understand. Usually few syllables indicate that a word is simple. Additionally, certain punctuation conventions should be followed. Avoid liberal use of paragraphing, with indentations and hyphens, in a long text. Separate paragraphs by blank lines and break up long text into several labels.

1.4.8. TESTS OF READABLE TEXT

Tests of text reading ability are widely used in the fields of education and linguistics. Two such tests which may be used to evaluate the reading levels of museum labels are Cloze Analysis and the Fog Index. The Cloze test is based on predictive indices. Words are deleted from the text at set intervals, such as every fifth word, and subjects who read this text have to insert appropriate words into the blanks. The Cloze test gives a reading score based on the percentage of correct insertions. A Cloze score between 57% and 61% indicates full comprehension of the material. The Fog Index is a much quicker and simpler measurement tool since it does not require human subjects. Its reading level assessment is based on a combined measure of sentence and word syllable lengths. The Fog Index was used to assess the reading level of a number of labels used in this study.

1.4.9. PRINT LEGIBILITY AND SIZE

The size of print used in a label depends on the importance of the information, the location of the label, the artifact size, the distance from the visitor, and the crowd flow. Such considerations result in the use of headlines, large letters at case and information focal points, medium letters to outline

important points, and smaller letters to impart less important information. Giving specific recommendations to designers, Weiner (1963) suggests the use of the following type point sizes:

- 72-90 points are good for distance viewing.
- 48-60 points for main and general text.
- 24-30 points (at least) for specimen captions.
- 10 points are permissible for very small objects, such as coins.

1.4.9.1. Label Type Face

Studies have been done on the influence of different type faces to enhance clarity and reading speed (see reports published by the Royal College of Art Research Unit). Simple conclusions can be drawn from such studies and used as broad guidelines for labelling in museums. One such conclusion is that the use of all capital letters reduces the reading speed, but is permissible for headings. Capital letters may also be used at the beginning of words, however, generally speaking, lowercase is best. The length of line should be relative to the typesize and should be 1 1/2 pica of width for each point of height.

1.4.10. COLOUR OF LABELS AND PRINT

Psychologists, Carmichael and W.F. Dearborn (see Royal Ontario Museum, 1976), have demonstrated that the colours of both print and background affect readability. They suggest the use of black ink on white or cream for minimum eye strain. They maintain that dark print on a light background does not strain the eyes, whereas, long reading against a dark or strongly coloured background is visually tiring. A neutrally toned background is a rest for the eyes, whether used in a whole display, or just for labels. Labels may blend with the background colour or be mounted on clear perspex. However such unobtrusive labels are best for the interested viewer.

1.5. THE IMPACT OF PREVIOUS RESEARCH ON THIS STUDY

This study inherits the problems and limitations of any attempt to develop diagnostic tools to measure human reactions and learning in a relatively unstructured environment. The inspiration to undertake such a task grew out of interest in other people's findings. Their example, their failures as well as their solutions and hypotheses, were instrumental in the formulation of this research. In particular, within the museum field, the idea of varying the display elements within a single exhibit case and testing the audience reaction to these changes, came from Parson's work in the Milwaukee Public Museum, and the taxonomy for delineating the relevant reactions came from H. Shettel's work. The Royal Ontario Museum's excellent handbook on exhibition design was the source-book for many of the attempts to portray "good design practice" throughout this study (including the layout of this thesis!).

CHAPTER 2

OVERVIEW OF

RESEARCH DESIGN

2.1. GENERAL RESEARCH DESIGN

This research was principally designed to test the following conditions:

1. LABELS- the use of pages of continuously paragraphed text *versus* text which has been broken up into spatially separated paragraphs with bold headings.

2. OBJECTS- thematically grouped objects *versus* an unstructured object arrangement.

3. GOALS- the presence of orientation statements and questions, placed at the beginning and end of the display case and aimed at stimulating the visitors' interest in the case, *versus* the absence of such external goals.

The term, "Principal" Display Treatment, has been used to distinguish between the different combinations of object and label layouts which were evaluated in this experiment, but not to differentiate between the presence or absence of goals. Within each of the three "Principal" Display Treatment samples, one-half of the visitors who were tested saw a particular treatment when goals were present, and the other half saw it when they were absent. In order to distinguish between the *With Goal* and *Without Goal* halves of the "Principal" Display Treatments, it is necessary to look at their subclassification, the six "Main" Display Treatment variations. (A full explanation of the experimental display treatments is provided in section 2.3, including the possible combinations of the conditions listed above which were excluded because they represented "bad" exhibit design practice.)

The primary measures used to evaluate the different display treatments, were the result of the unobtrusive observation of every visitor to the gallery,

and the result of a questionnaire which was administered to every visitor who fulfilled the following three criteria: 1) they must have appeared to be fourteen years of age or older, 2) they must have spent a minimum of 45 seconds looking at the experimental display case, and 3) they must have at least glanced at all four of the main sections (the 'Ages') within the display case.

The observation studies yielded cumulative results of the *Initial Attraction*, of all visitors present, to the display case (i.e. the proportion of visitors who went directly to the experimental display case after entering the gallery), and their *Interest Retention* (as measured by the mean amount of time visitors looked at the experimental display case). The questionnaire yielded additional information about the selected visitors' *Knowledge Gain* (the amount of learning which took place, based on the results of a short test), and the visitors' viewpoints as expressed through their *Written Comments*. These results were assessed by using frequencies, means, chi square, analysis of variance, and multiple regression analysis techniques. Comparisons were also made with the survey data from other museums.

These display and evaluative techniques were implemented over a five month period using a specifically designed upright museum display case which contained prehistoric artifacts, photographs, and drawings. A total of six hundred voluntary visitors to the prehistoric archaeology gallery of the National Museum of Antiquities of Scotland, in Edinburgh, were asked to fill out a questionnaire and test of their exhibit knowledge, after they had seen a "Main" Display Treatment. Testing averaged ten days for each of the six Main Display Treatments and included both weekday and weekend visitors. In addition the initial movement patterns after entering the gallery, and the length of time spent viewing the experimental display case, were recorded for all visitors present.

Due to the "real life" conditions in which this study was conducted, as opposed to a strictly controlled and set up laboratory or matched sample experiment, it was impossible to rigidly adhere to a strict experimental design. Instead a quasi-experimental design was utilized. Unlike a strict experiment, in which the membership of different treatment groups is normally determined by randomisation and matching, in this quasi-experimental format, both the treatment groups and the control groups contained unmatched samples who operated under unmatched conditions. (Since this was not a rigorously designed experiment, all terms such as "control group" should be interpreted accordingly.)

Because the design was not a true experiment, it was possible for both testing conditions and the characteristics of sample members to vary between treatments. To assess and control for this variation, a great deal of supplementary information was recorded for visitor movement patterns; environmental conditions; demographic characteristics; and avowed prior knowledge of, and interest in, the display topic. This supplementary information was extensively analysed in order to test for any bias in the results which might have been due to differing treatment conditions or populations. Pilot studies and control groups were used extensively to refine the test procedures and to ensure that the final results would be valid. Such studies included pre-test/ post-test, 'test yourself', and 'asked to study' control groups; and the evaluation of the old display, new pilot display, new small label display, and a re-test of the first display variation at the end of the evaluation period.

2.2. BACKGROUND CONTEXT TO THE STUDY

2.2.1. Initial Interviews with Staff from Several Museums which Established the Need for Research

Informal interviews were carried out with design, curatorial, and educational (where they existed) staff from several museums in Britain and America in order to establish a relevant focus for this study. It quickly became apparent that many museum staff members were unwilling to change their displays (even the Victorian ones!) because they were not convinced that modern display techniques would necessarily be an improvement. Many curators expressed their frustration in not knowing how to put across their enthusiasm for their subjects to the museum visitors, in a static display. Many noted the lack of any general guidelines for improving museum display formats (as opposed to specific design variables such as size of print). This study was undertaken in an attempt to provide a few such guidelines.

2.2.2. Contact with N.M.A.S. Staff

The National Museum of Antiquities of Scotland was chosen to host this study because it was beginning the gradual re-display of its entire archaeology collection. Thus the opportunity existed to implement the successful display solutions on a large scale. In order to ensure that the focus of this study would remain relevant to members of the museum staff, various interviews were conducted to ascertain how they viewed the overall display policy of their museum and its relationship to the public.

2.2.3. Location of the Study and Description of the Museum Collection

This study took place in the first floor gallery of the National Museum of Antiquities of Scotland (N.M.A.S.), Queen Street, Edinburgh. That gallery contains an exceptionally fine and comprehensive collection of Scottish archaeology ranging from the first appearance of Man in Scotland (circa

6500-6000 BC) through to the Viking incursions and settlements (9th-12th centuries AD). The techniques used to display the collection also varied widely in their life range, from recent displays through to those which have remained unchanged for almost a hundred years (and still retain their handwritten labels!) In addition, the unvarying location of many of the displays makes it possible for the 1898 catalogue to remain a valid reference for visitors seeking specific information!

The nature and size of the museum's collection needs to be fully borne in mind because it has largely governed the display policy in the past. Since the end of the nineteenth century the museum has occupied one-half of the Scottish National Portrait Gallery's "many-windowed" Gothic revival building. The archaeological material is contained in two galleries which are far too small to do them justice, much less provide space for recent acquisitions. Superimposed on this major space restraint has been a display policy governed by the belief that visitors are primarily interested in seeing the archaeological finds from their own areas. Furthermore, the staff's avowed desire to represent "the forefront of knowledge about the artifacts" in their displays has been tempered by the expectation of a new building and new display cases.

Space restrictions have encouraged the use of extremely dense displays both in and under the numerous tall wooden display cases. Considerable repetition of artifacts (for example, 33 similar-looking stone axe heads in a case) was also much in evidence due to the belief that visitors were primarily interested in seeing the specific archaeological finds from their home localities. However, there was a growing feeling among many of the museum staff that the appreciation of such a display was largely confined to archaeology students or to visitors who came specifically to see one or two items (even if they do look like most of the other axe heads on view). The then director of the museum, Mr. R.B.K. Stevenson, expressed the idea that rows of objects on view

were more suitable for the tertiary sections of a gallery rather than for primary exhibition space.

2.2.4. Display Case Vehicle

A display case was chosen as the vehicle for this study because, as in many museums, cases provide the bulk of the museum's interaction with its public, and apparently will continue to do so for the foreseeable future. In this experiment an in-depth study was made of the introductory case since it offered a synopsis of the range of material and content on show in the gallery. The 12 feet long wooden display case, typical of the upright display cases in the gallery, was located to the left, and in front of, the only public doorway into the gallery. A large sign beside it proclaimed that it was the "Introductory Case". The case was large enough to allow a wide variety of objects to be on show but small enough that one observer could reasonably time all of the visitors looking at the case at any one time. It was also small enough to allow an acceptable transfer rate of between 3-4 days to make significant changes between experimental display treatments.

2.2.5. Time Frame For This Study

Initial consultations with staff at the National Museum of Antiquities of Scotland began in October of 1977, as did observations of traffic flow patterns on the floor of the gallery. The preliminary pilot studies began in December. Over the next few months refinements to the questionnaire were explored together with new design structures and suitable label content for the experimental display case. Observation and testing of the resultant exhibit variables took place every day of the week, during the hours that the museum was open, from March, with only occasional interruptions, until the middle of July 1978.

2.2.6. Population Sampled:

This study was partially instigated to provide information about the visitors to the museum, as a whole, in their approaches to, and interest in, a designated potential learning situation. Consequently, a record was kept of the initial reactions of all casual museum visitors to the experimental display treatments. These general observations were recorded for all of the 2,713 visitors present in the gallery during the principal portion of the experiment, i.e. when the principal 600 visitor sample was tested.

Information about the visitors' recall of the designated potential learning situation was confined to those visitors who qualified for, and completed, a questionnaire. The number of visitors to be tested with a written questionnaire was determined before the principal experiment began. The first 200 visitors to the museum who were over the age of fourteen, who viewed the overall introductory display case (all four parts) for more than 45 seconds, and who agreed to participate in the experiment, comprised the tested sample population for each of the three 'principal' display variations. Within each principal 200 visitor sample, 100 visitors saw that display variation when goal statements were present, and 100 saw it when goals were absent. These 100 visitor groups made up the tested samples for the six 'main' display treatments. After 200 visitors had seen one of the principal display variations and had been asked to complete a questionnaire, that display variation was removed and another installed in its place.

For the three principal display variations, 600 visitors were engaged in a written test, and, as were all of the visitors, were observed for their initial traffic patterns and length of attention to the experimental display case. If, however, control groups, secondary experimental groups, and pilot study groups are included in the head count, approximately 1000 visitors to the museum participated in the written portion of this experiment.

2.2.7. Restrictions on the Choice of Design Techniques

After much deliberation and debate with relevant members of the museum staff (i.e. the curator and designer who were working with the study) about the most appropriate methods of organizing the artifacts and labels, the three Principal experimental design "treatments" were decided upon. These display treatments were selected because they were suitable for testing purposes as well as being representative of both existing, and potentially acceptable, display techniques.

Several practical constraints influenced the choice and presentation of these "treatments". Necessary materials had to be readily available to a museum design department as well as being inexpensive. Only a few objects could be selected which were either too large or too fragile to be moved about. Labels had to be written so that each paragraph could exist autonomously and would relate to a particular theme. Specialist reconstructions such as dioramas, feathered arrows, and axe handles were beyond both budget and time restrictions imposed by the museum. The techniques utilized also had to be consistent with the overall tone and sensibility of the museum and fit into its standard sized upright Victorian display case.

2.2.8. General Arrangement Of Objects And Labels

The introductory case was organized to provide a synopsis of prehistoric Scotland, i.e. Scottish archaeology, and hence history, from the arrival of the first men (c. 6000BC) until approximately 1100AD when written accounts begin to play a large role in interpreting the way of life in Scotland. Following established archaeological procedure, this time span was divided into four "Ages", i.e. Stone Age, Bronze Age, Early Iron Age, and Late Iron Age. These divisions were shown on four vertical panels within the case, each nearly three feet wide, presenting representative artifacts and photographs for each "Age". The objects within the experimental introductory display case were further

selected to all fit into four "Thematic" categories: tools, weapons and defended settlements, ornaments, and burials, examples of which could be displayed within each "Age".

From a purely design orientated point-of-view, an equally valid arrangement of the display case would have been to position the "Themes" in the vertical panels and the "Ages" in the horizontal bands. In fact that type of arrangement is much favoured by archaeologists since it gives prime importance to the concrete, known, objects rather than to the controversial, speculative, and somewhat arbitrary, "Age" categories. However there were two very good reasons for not using that arrrangement. In the first place, it would not have allowed the observer to easily monitor visitor learning sequences by noting the subjects' movement patterns between the vertical panels. For example, viewing Tools before Ornaments would not suggest an irregular learning sequence but viewing the Late Iron Age before the Early Iron Age would. Secondly, that arrangement would not lend itself to the use of primary labels containing general observations about lifestyles and changes across a time period. Such socially orientated label content may often be of interest to the visitor, help to establish a context for subsequent information, and help in establishing rapport.

New labels were written for the case and sent to Dr. Clarke, the museum specialist in prehistoric archaeology, who red-pencilled them to ensure that the information was not simplified beyond accuracy. The labels were colour coded depending on the information they contained. Cream-coloured backgrounds were selected to present the primary, descriptive information; gold-coloured cards displayed secondary information (mainly dealing with archaeological sources for the periods); and green labels were used to identify specific objects (because they blended in with the green felt background of the display case, thereby causing minimal visual distraction from the artifacts themselves).

In designing the interior of the chosen display case, as well as the presentation of the chosen label, object, and goal layouts, much thought went into creating an interesting, yet flexible, display. Throughout the experiment the actual objects and the content of the labels in the experimental display case remained invariable, as did the background colour and the other design elements in the case which were not being studied.

2.3. EXPERIMENTAL TREATMENTS:

The display techniques evaluated in this study involved two opposing label and object treatments and the inclusion or exclusion of goal statements directed at the visitors. Within the single museum exhibit case which was the focus of this study, the following variations were successively interchanged:

2.3.1. Label Variations:

B= single 'blocks' of primary information containing contiguous paragraphs, similar to a page in a book.

M= a 'multiple' series of bold headings and physically separated paragraphs of primary information relating to a particular topic, somewhat resembling a newspaper or outline format.

2.3.2. Object Variations:

U= an 'unstructured' or random arrangement of objects.

T= a 'thematic' grouping of objects into horizontal bands of tools, weapons, ornaments, and burials.

In a single experimental display 'treatment', one of the label variations (**B**lock or **M**ultiple) would be combined with one of the object variations (**U**nstructured or **T**hematic). The possible resultant combinations, termed "Principal Display Treatments", which were evaluated in the experiment were:

- **MTs** =**M**ultiple labels with bold headings and **T**hematically grouped objects.
- **BTs** =a single label **B**lock and **T**hematically grouped objects.
- **BUs** =a single label **B**lock and an **U**nstructured, or random, object arrangement.

The fourth possible treatment, **MUs**, multiple labels with unstructured objects, was not carried out or evaluated in this experiment. It was felt to represent an illogical display treatment which was unlikely to be used in a museum setting and was consequentially irrelevant as a practical design alternative.

2.3.3. Goal Variations:

A consistent set of avowed aims, objectives, and interrogative statements, otherwise known as orientating statements or 'goals', were present for half of the subjects (i.e. 100 visitors) who were evaluated in each of the three principal experimental display treatments, e.g. **BUs**, **BTs**, and **MTs**. The subclassification "Main Display Treatments" has been employed to indicate the presence or absence of Goals as well as to distinguish between the different object and label formats.

The presence of the letter "**G**" at the end of an experimental treatment designation, such as MT**G**, indicates that **G**oals were present during that treatment. Conversely, the absence of the letter "**G**" terminating a treatment designation, such as **BU**, indicates that Goal statements were also absent from the display. Goal statements were only present during three of the "Main" experimental variations and were not present during any of the pilot or control treatments.

- "Without Goals" =**MT, BT, BU**
- "With Goals =**MTG, BTG, BUG**

There follows a summary of the treatment designations. The experimental variations in the display case fall into several different categories. First in importance are the three "Principal" experimental treatments, **MTs** (=MT+MTG), **BTs** (=BT+BTG), and **BUs** (=BU+BUG). These "Principal" display treatments reflect differences in Object and Label layout, but not in Goals. Two hundred visitors were tested in each of these "Principal" treatments. The three "Principal" treatments were subdivided to form the six "Main" experimental variations, i.e. **MT**, **MTG**, **BT**, **BTG**, **BU**, **BUG**. These treatment designations distinguish between Object, Label, and Goal arrangements. There were one hundred visitors tested in each of the "Main" experimental treatments. The "Main" treatments may be grouped into two categories, **With Goals** (MTG, BTG, BUG), and **Without Goals** (MT, BT, BU). Both contain three hundred in their tested visitor sample populations.

It is important to understand what is meant by these treatment designations since they will all be used in the analysis of the study results. It is also important to realize that they collectively relate to the same visitor population. In all of the above treatment groups, the first letter of the designation refers to the Label variation, either an **M** for multiple labels, or a **B** for a label block; the second letter to the Object variation, i.e. **T** for thematic object groups, or **U** for unstructured object groups; and the presence of a third letter, **G**, indicates that Goal statements were also present in that variation.

2.3.4. Pilot and Control Groups

Only two additional experimental treatments utilized the same test, observation factors, and visitor sample size as did the "Main" display treatments. These were the Control treatments, ***MTa*** and ***RMT***. ***MTa*** was a replica of the "Main" display treatment, **MT**, except that it had smaller headings. Therefore, by comparing the results from these two treatments, it should be possible to assess the impact of small versus bold headings. The same

experimental treatment, **MT**, was tested both at the beginning and at the end of the "Main" experimental treatment block, several months apart, in order to assess the likelihood of potential misleading results due to seasonal differences in the visiting population. The second time it was tested, it was re-designated as the Control treatment, ***RMT***.

Other Pilot and Control treatments, some of which used different evaluative tools from those used in the main experiment, are discussed below. It is important to understand the distinction between the terms "Pilot" and "Control". Through the analysis of the results of a succession of "Pilot" studies, sign, observation, and test formats and questions were changed and refined. "Control" studies and treatments, on the other hand, refer to a fixed group or treatment which was specifically designed to be analysed in comparison with the "Main" treatment results.

Much effort during the pilot-testing phase went into re-wording the questions and arranging the choices to minimize the number of correct guesses which took place. Various pilot questionnaire formats were tried out on visitors during the ***Pilot Questionnaire Study***, before the questionnaire was ready to be used with the "Main" Experimental Treatments. It was not possible however, given the space restraints imposed upon question format and length by the decision to limit the questionnaire to two pages in length, to reduce the average number of correct responses to the pilot questionnaires below 16%, without increasing the complexity and narrowing the focus of the questions themselves. The pilot questionnaires were administered to all visitors who entered the gallery, before they had a chance to look at any of the display cases.

Additionally, interviews were conducted with a small number of visitors who had seen and been tested on one of the display treatments, to ensure that

specific questions were not misleading them, and that their responses did, in fact, reflect their knowledge gain (see the next section for an explanation of these terms), written comment opinions, and background characteristics. The final questionnaire version was used for all tested groups, apart from the ***Pilot Questionnaire*** and ***Alleval Rating*** groups.

Also during the ***Pilot Questionnaire Study*** phase, before the start of the Main Experimental Treatments, observational data alone was collected for a variety of sign characteristics, such as: formats, statements, print sizes, background colours, and locations. Changes were made to these characteristics, and the results were evaluated for their ability to direct the maximum number of visitors to look at the experimental display case. The final sign, used for the Main Display Treatments, resulted from the assessment of this Pilot observational data.

Observational studies were also conducted of the *Initial Attraction* value (see the discussion of this term in the following section) of the old display and of the new display containing difficult text. Since the label information in both of these displays was different to that used in the Principal display treatments, written questionnaires were not administered. References to the ***Yellow Case Control*** refer to the old display treatment, and to the ***Green Pilot Case*** (which was used as a Control for observational studies) refer to the use of the new display background colour and interior fittings, but with slightly more objects and more difficult text.

The Control Groups established values for comparison with the Main Display Treatments. As in the Main Display Treatment experiments, all of the subjects who took part in the Control experiments were voluntary visitors to the museum, apart one group of high school students who were asked to come to the museum to participate in an experiment, the ***Firhill*** group.

One of the earliest Control experiments involved a group of visitors to the gallery who were asked to complete the questionnaire before they had seen any of the display cases on the first floor, the ***Pre-Test*** group. All of the teenage and adult visitors who entered the gallery during this Control experiment were asked to complete the questionnaire in this way. By asking for the cooperation of all of the visitors to the gallery during this experiment, it was hoped that the resultant sample would be reasonably representative of the visitors to the gallery as a whole. Since this was not a primary experiment, the available time was limited and the resultant sample size was small, as were most of the samples in the other Control experiments. In all, twenty-two visitors completed the ***Pre-Test*** and were then requested to look at the experimental display case before filling out the questionnaire for the second time that day (the ***Post-Test***). Therefore the same twenty-two visitors were in the ***Pre-Test*** and ***Post-Test*** groups. The ***Pre-Test*** group established the average correct score that visitors knew or could guess before they saw the experimental display case. The ***Post-Test*** group, on the other hand, demonstrated how much visitors could score who had taken the pre-test, seen the display case, and then taken the same test again.

The only Control group, other than the ***Pre-Test*** group, which completed the questionnaire *before* they looked at the display case, was the ***Pilot Questionnaire*** group, otherwise known as the ***Pilot (unseen)*** group. The results for this group were taken from questionnaires which were very similar to the final questionnaire version. All visitors who entered the gallery during this experiment were asked to complete a questionnaire. Therefore this Control group may be used to compare the demographic characteristics of "all gallery visitors", as well as their reasonably comparable test scores before seeing a display case, with the characteristics of those who were tested in the Main Display treatments.

The results from observation studies of preliminary display treatments may also be used to compare with the observation results of the Main Display Treatments. Several of these comparative display treatments were: the ***Yellow Case Control*** (the original display), the ***Green Pilot Case Control*** (difficult labels), the ***MTa Control*** (small headings), and the ***Alleval Rating*** treatment (discussed below). Photographs of all of these Control display case treatments may be seen in the Appendices.

In several other Control experiments, visitors were specifically directed to look at the case and to complete a questionnaire. These Control groups helped to establish whether or not the answers to the test portion of the questionnaire could be found in the display case. The ***Ask Study*** group were approached and asked to look at the display case for as long as they wanted to, knowing that they would be asked to take a test on the information in the display afterwards. Similarly, the ***Firhill*** study recorded the results of a small number of male high school students who were asked to look at a display treatment knowing they would then be tested on what they had seen. These schoolboys were specifically brought into the museum in order to gauge the ability of boys in the fifth form to complete the test after viewing the display case (see the ***Firhill*** results). The ***Test Self*** group were visitors who voluntarily picked up, completed as they were viewing, and returned, the questionnaire-test, which had been appropriately displayed to elicit this response (see the photograph of the "Test Yourself Box", in Appendix B.V). These visitors were directed to find the answers to the questions while they were standing in front of the display case, in order to discover whether or not it was possible for museum visitors to obtain high scores on the test. The results from all of the Control groups mentioned may be found in Appendix G.

In the ***Alleval Rating*** study, visitors were given the opportunity of seeing several of the main experimental variations side by side in the display case.

The visitors were asked to complete a specifically designed rating sheet which evaluated certain aspects of the three principal display treatments. The ***Alleval*** display case was arranged so that sections of **MT, BT** and **BU** were shown concurrently on separate panels. A photograph of this display arrangement may be seen in Appendix B.IX, and the results from the rating sheet are reported in Appendix F. This type of rating sheet may represent a simple way for museum personnel to obtain feedback on their exhibits, since they are easy to administer and mark, and they take up little of the visitors' time.

2.3.4.1. CONCURRENT EXPERIMENTS:

Several concurrent studies were set up to provide tangential information on a number of areas. Results from them may be found in the Appendices and in the discussion of statistical analysis in Chapter 6. Some of these secondary experiments were: 1)the predictive usefulness of time studies and interaction counts to determine how much visitors' learned from the display, 2) visitors' initial movement patterns in a non-formal learning environment, and 3) the influence of avowed knowledge and interest in the subject matter on subsequent attention and learning. Other tangential studies yielded cursory conclusions regarding the impact of large versus small headings, difficult versus easier text, the old display versus the new, a directional sign, and a 'test yourself' sign. These findings are discussed under Control and Pilot Study results in Chapter 6.

2.4. EVALUATION MEASURES OF DISPLAY EFFECTIVENESS

In order to assess their 'effectiveness' the experimental display case treatments were evaluated in these four areas:

1. INITIAL ATTRACTION- measured by the proportion of

visitors who viewed an experimental display treatment before looking at any other display cases (the total number of visitors present divided by the number who initially went straight to the experimental display case).

2. INTEREST RETENTION- a measurement of the average length of time that all visitors present looked at an experimental treatment display (the total viewing time divided by the total number of viewers present).

3. KNOWLEDGE GAIN- measured by test scores.

4. WRITTEN COMMENTS- broadly recorded as the resultant total number of positive respondents and comments after all negative respondents and comments have been subtracted. Additional analysis of comment categories has been done by comparing the number of positive and negative comments per display treatment group for each comment category.

The *Initial Attraction* and the *Interest Retention* measures both record the observed actions of all of the visitors who were present while an experimental display treatment was being evaluated. The *Knowledge Gain* and the *Written Comment* assessments, however, only apply to the museum visitors who were asked to complete a written test and questionnaire.

The measurement tools used to evaluate the communication effectiveness of the relevant display treatments were: observation of initial visitor movement patterns, display case viewing times, test results, and written comments. (For a full discussion of the evaluation measurement tools used in this experiment, see Chapter 4 .) Observation records were kept for all of the casual visitors to the Gallery. The total number of visitors, their initial movement patterns, and the size and composition of their groups, and the length of time they looked at the experimental display case were noted. Additional information was recorded for those visitors who were tested as part of the "Primary" treatment samples. This included the size and composition of their groups, their interactions, and their direction of approach to the case and sequence of viewing it. Several general observations were also recorded at the beginning of each day of

testing regarding the weather, the noise and crowding conditions in the Gallery, the calendar date, and the day of the week.

A questionnaire was developed to test the visitors' retention of specific information found in the experimental display case and to elicit demographic material and written comments. After the visitor had studied the case for 45 seconds or longer, and had moved away from the vicinity of the experimental display case, s/he was asked to fill in the questionnaire at a table behind the case and to drop the completed questionnaire in a closed, marked box beside the table. (for a discussion of the results of these evaluation measures, see Chapter 6)

The *Knowledge Gain* measurement was derived from the results of a test. This test formed the larger part of a two page questionnaire. In format, the questionnaire began with an introductory statement soliciting participation, closely followed by the test- eighteen knowledge questions based on information which could be found in all of the display treatments. The questionnaire continued with a number of demographic and background questions and a request for free formatted *Written Comments*. The concluding statement acknowledged appreciation for participation.

In calculating the test scores, each of the eighteen items were given equal weight, with the obvious result that "18" was the highest score possible.

CHAPTER 3

ORGANIZATION OF

TESTING ENVIRONMENT:

DISPLAY CASE, OBJECTS, LABELS, GOALS, AND SIGN

3.1. DISPLAY CASE:

3.1.1. Selection of the Display Case Format– Internal Validity:

There are a number of advantages of, and precedents for, testing a succession of design treatments in a single display case, as opposed to testing several simultaneously shown display cases or areas. The single display case has the advantage of being cheaper and less time consuming to design, its display treatment variables are always seen in the same physical location, and it is easier for one observer to monitor visitor movements around one display case than around several. Most importantly, this successive treatment allows alternative methods to be tested using the same objects, text, background, format location, and test instruments, so that exact parallels can be drawn. Since the fundamental aim of this study was to evaluate alternative object, label, and goal treatments, the single display case provided an ideal opportunity to do so without bias from differing locations and content. In the Milwaukee Public Museum, both Abler (1968) and Parsons (1965) evaluated visitor reactions to successive display treatment variations in one display case. Such studies can produce implications for the wider environment, both within the host museum and others.

Three alternative display formats were rejected: simultaneous testing within an exhibition hall, conducting a survey of visitor reactions within the entire

museum, and using photographic mock-ups of displays tested in a laboratory situation. Available space, personnel, and funding were inadequate for the major changes necessary to test specific design variables in an exhibition. Likewise the diagnostic tools and personnel were inadequate for a survey of an entire museum. Both situations offered additional disadvantages. One could not control for chosen variables such as: object, label, and display locations, or test comparable material without blatant replication.

Mock-up displays may be either 'life-size' photographic blow-ups of a display, or photographs reduced in size. These could be set-up anywhere and questions based on them could be put to a range of subjects, including museum visitors. Whereas measures of potential learning could be based on these mock-ups, they could provide no measure of the attraction and holding power of the exhibit. In addition no comparisons between mock-ups and 'real' exhibit testing were available to provide a realistic assessment of their predictive usefulness. The advantage of a mock-up was the ease with which it could be modified and the subject be presented with a variety of situations in the course of testing. Mock-up displays in a artificial environment introduce the greater possibility of tester bias and influence. Although a previous study has suggested that the results in some situations may be comparable to those obtained in a museum (Shettel, 1973), the results are much less likely to be viewed seriously, and acted upon, by the museum profession.

3.1.2. Location of the Display Case within the Museum:

3.1.2.1. Choice of floor-

Several conditions within the museum dictated the choice of floor, or gallery, for the study. (Broadly speaking, the layout of the three floors at the time of the study corresponded to single halls or galleries, each representing an autonomous area of the museum's permanent collection.) The ground floor

was unavailable due to alterations and closure for part of the evaluation period. On the second floor, a comparatively small collection of objects were more openly displayed than on the other two floors. This floor only contained Roman influenced objects from the relatively short timespan of the Roman occupation of parts of Scotland. Much smaller than the other two floors, it provided no "hidden" areas where visitors could sit down at a table to complete the questionnaires, or where an observer could lurk unobtrusively. Another reason for avoiding this floor is that museum studies (O'Hare, 1974) have suggested that, for a general collection, the higher up the floor and the further it is away from the museum entrance, the fewer are the numbers of visitors who are likely to reach it.

Instead the first floor was chosen. There was great variety and challenge in the objects and display formats represented there. The display cases on this floor were exceptionally varied, some unchanged from their hand labelled, Victorian state, some full of row upon row of similar-looking dull objects, and a few with up-to-date photographs and labels relating the objects to the geological and environmental context in which they would have been used. Therefore this floor provided a context in which a "different-looking" display case would not appear out of place. Areas were also available where an observer and subjects could write without attracting much notice.

3.1.2.2. Near the entrance to the first floor-

A display case was chosen for this study near the entrance to the first floor gallery. Its location maximized the chance to attract the largest possible viewing population, before their attention was caught by other areas. Since the majority of visitors viewed the display case soon after entering the floor, it represented the best choice for test question validity. The amount of information which they were likely to have gained from other sources in the gallery was thereby reduced.

This proximity to the entrance was also necessary in order to observe the visitors' initial traffic patterns. The layout of the gallery entrance allowed the obviousness of an observer/tester to be minimized. The observer often went unnoticed at a desk behind the door or was assumed to be part of the general museum staff, i.e. attendant or information personnel.

3.1.3. Type of Display Case Chosen:

As with most of the display cases in the museum, the one chosen for this study was of an upright, old-fashioned design, with storage facilities in its base concealed behind dark coloured wooden panels. Its approximate dimensions were 12 feet long by 6 feet high by 3 feet wide, within which the top 4 feet were glazed on three sides. The base and the back of the case were made-up of wooden panels. The top and sides were glass set into wood. Four glass doors, each approximately 4 feet high and 3 feet wide and surrounded by wood, made up the front of the case. The case had no internal lighting or temperature control. The front of the case was divided into four sections by the wood surrounding the glass doors and the support bars in between them. This division into four, three foot sections, was only on the surface of the case and did not extend into the interior.

3.1.3.1. Reasons for the selection of the Introductory Case-

For a number of reasons, the Introductory Case was chosen for this experiment. It was representative of the entire collection since it presented both an overview and a summary of the types of objects displayed on the floor. It maximized interest potential since the case, by definition, contained a wide range of items which appealed to a broad interest spectrum. As a primary orientation case, it provided the opportunity to organize part of the collection in the gallery into reference categories. Thus, not only were specific items in the case put into context, but focal points and background material were provided which could potentially lead to an understanding of the

relationships between other objects throughout the gallery. It could also logically use display techniques which differed from those on the rest of the floor since it was a different type of case. There was also the enhanced likelihood that many visitors would spend a greater amount of time and attention viewing this case due to the absence of other orientation materials and signs.

3.1.4. Display Case Constraints:

There were a number of disadvantages to the display case chosen. Its location on the left, out of the way, not in an obviously apparent position, was a potential disadvantage. Museum visitor traffic studies in America (Melton, Psychological Bulletin 1933) have suggested that most people, when given a choice, will turn to the right after entering an exhibition. Consequently, a position on the left was not likely to be advantageous. (Since there were more visitors from North America than from Britain in this study, North American traffic patterns were deemed to be the most significant in posing locational advantages and disadvantages.) Its location also forced it to compete with the attractive display of St. Ninian's Treasure nearby. Furthermore, the attendants often stood by the door, near the case, which put off those visitors who disliked being watched.

The case was very high. Its top display space was not particularly effective since many things go unnoticed above eye-level height. An effective layout was also potentially hampered by the physical division of the case due to its three wooden supports located between the glass doors on the front of the case. These supports ran vertically from the top of the case down to the top of the base unit, at three feet intervals, dividing the display area visually into four units.

The case was also fairly deep. This had the disadvantage that it did not

allow visitors to closely view the objects at the back of the case, or their adjacent labels. However it did allow a variety of depths to be created by using background panels, and reasonably large objects could be accommodated.

Unfortunately the case had no internal lighting. Instead, it relied on artificial lighting from the ceiling and natural lighting from the large windows on three sides of the floor. Consequently there were problems with glare and reflections from windows and ceramic tile wall sections. Additionally, there was the possibility of distractions seen through the glass side panels.

Therefore disadvantages occurred because of the display case size, depth, layout, location, lack of internal lighting, and potential for glare and distractions. Such disadvantages, however, may have equally applied to most of the other display cases in the gallery.

3.1.5. Layout of the General Case Design:

The general case design reflected the decision to basically strive for "good practice" in all six of the main display treatments. Therefore this study was not trying to maximize the variability between treatments by testing at least one "bad" design. Instead, the difference between treatments is in the context of generally good design practice.

The Introductory Case gave an overview of the timespan and range of objects found in the Scottish archaeology section of the museum. It provided a synopsis of Scottish archaeology, and hence prehistory, from the arrival of the first human visitors (circa 6000 B.C.) until circa 1100 A.D., when written accounts begin to play a large role in interpreting the life of Scotlands' people.

3.1.5.1. Division into 'AGES'–

Following established archaeological procedure, the timespan covered by this case was divided into four 'Ages': Stone, Bronze, Early Iron, and Late Iron Ages. Objects representative of these 'Ages' were arranged vertically on one of four separate panels for each 'Age'. These panels corresponded in height and breadth to the dimensions of the glass doors fronting the case. The Stone Age panel was displayed in the far right hand section of the case and the other Ages followed in chronological sequence, finishing with the Late Iron Age panel on the far left.

- Stone Age (6000–2000 B.C.) subdivided into Hunters and Farmers.
- Bronze Age (2000–600 B.C.)
- Early Iron Age (600 B.C.–300 A.D.)
- Late Iron Age (300–1100 A.D.)

3.1.5.2. Division into 'THEMES'–

The artifacts and photographs were also selected to fit into one of four 'Thematic' categories: Tools, Weapons, Ornaments, and Burials. Objects representative of these themes could be grouped in horizontal bands across the case. The same thematic groupings could then be displayed within each 'Age', starting with Tools at the top and going down to Burials at the bottom.

- Tools
- Weapons and Defended Settlements
- Ornaments
- Burials

All items within the display case fitted into both an 'Age' and a 'Theme' category.

A single background colour was used in the display for several reasons. It

minimized the influence of locational differences when testing objects and labels in different places in subsequent design variations, since individual colours have varied attraction values for different people (Dandridge, 1966). The use of several colours was also unsuitable due to their potential for creating inequalities in test variations or for over-emphasising particular areas. For example, using colour banding to distinguish thematic groups was undesirable because objects in one of the test variations (**BU** -block, unstructured) were not grouped into thematic categories; and using colours to differentiate between the 'Ages' was also undesirable because the colours could detract from the continuity of the themes.

A lime green background colour was chosen to emphasize dull objects such as stone and iron fragments and to somewhat de-emphasize the bright attraction of gold and silver items which previously received inordinate emphasis. The designer's brief emphasized the desire to give equal weight to all objects. To accomplish this aim, it was necessary to try to make the Stone Age section appear as attractive as possible since it was frequently overlooked by visitors in the pilot studies.

Visual complexity was minimized by using various design elements to lead the eye on, such as, similar size placement, interlinking shapes and backgrounds, as well as a simple colour scheme. Distractions were also minimized. Photographs and drawings were placed as far away from the cream coloured labels as possible to minimize distractions from similarly coloured backgrounds. Full height, width, and depth side panels were introduced to all the principal display treatments, in matching lime green felt. These were introduced to hold goal labels and other orientation material, to focus attention within the case, and to reduce reflections and distractions. The use of variable depth of field panels between the four vertical 'Age' divisions in the case, dispensed with the need for zigzag strings or other vertical separating devices.

3.2. OBJECTS:

3.2.1. Selection of Specific Objects:

To be selected for this display, objects needed to fulfill a number of predetermined criteria. An object had to fit into both 'Age' and 'Theme' categories and be representative of its type. To qualify for inclusion in the Introductory Case, it also had to be representative of areas of emphasis within the museum's collection. It also had to be available for use. Some objects were on loan to another museum and others were not available in this museum's collection. Objects were also selected for their human interest value, because they were interesting: intrinsically, visually, in context, or by connotation. They were also chosen because they illustrated a point, told a story, etc. Some objects were rejected because they would take up too much space or because they were too heavy to be mounted easily on a vertical surface, ex. a grinding stone. Limitations were also created by the number of objects which could be displayed in the case.

Objects which required a great deal of explanatory labelling were rejected in favour of ones which could be made readily intelligible through reconstructions or illustrations. Non-objects such as drawings, diagrams, or photographs, were used instead of long labels to show items in context or as they might have appeared in use. Photographs were used for general rapport and human interest. They were also used to establish the context for objects and ideas and to show human settlement patterns and environments. Details of objects were shown side by side using photographic blow-ups, as with the gold lunula and the jet spacerplate necklace, to highlight similarities. The only time a photograph was used in lieu of an object was when an object was unavailable or was prohibitively large.

3.2.2. Object Limitations:

A number of factors had to be overcome, or minimized, which detracted from the objects selected for the display. Most of these prehistoric objects were dull and not immediately attractive. Many items were composed of broken fragments or were incomplete. As a whole they presented a very broad mixture of sizes and shapes, from large objects which took up a great deal of space, down to items so small that they could easily be overlooked. Additionally fixed perspex holders detracted from certain objects. However as far as possible, the distraction of holders and broken fragments was minimized by concealing them.

Additional problems occurred because very fragile objects had to be mounted on movable backgrounds so that the actual object need not be handled when changing the display. Also certain items and shapes, such as necklaces and heavy, rounded, axeheads, were difficult to mount on a vertical surface. Large heavy pots had to be placed on a flat surface. In addition several important objects planned for the permanent display were sent on loan to the Aberdeen Museum when testing began. They had to be replaced by illustrations, ex. the jet spacerplate necklace, the Huntarian Brooch, a Pictish chain. They were not reinstated in the case until the end of the experiment.

3.2.3. Object Placement:

The following criteria, based on subjective views of good design practice, governed the placement of objects:

1. SHOWN IN USE– As far as possible, objects were shown as they would have been used, ex: a spindle whorl was inserted onto a rod which was hanging from threads of fleece, small pins were inserted into a cloth, a scraper and an awl were positioned as if they were working leather, limpet hammers were positioned as if they were striking limpet shells, a necklace was spread out as if were around someone's neck, composite tools were inserted in a shaft, and a weaving comb and weight were shown in use with a reconstructed loom section.

2. ILLUSTRATED USE- Drawings were also used to demonstrate an item's function, ex: drawings of an axehead and of a plough share in place, each beside the actual item, which was angled to match its position in the drawing.

3. IMPLIED CONTEXT- Objects found in the same context were placed together, for instance, the Viking burial objects were displayed above a photograph of a Viking burial. The use of the funeral urns in the case was implied by placing them in recesses to impart the feeling that they were hidden or buried, ex: in a shelf in the Stone Age, in a cist in the Bronze Age, etc. These burial recesses not only suggested that the pots were covered over, but also hid the gaps in the pots! Relevant photographs were used to show the pots in a typical burial arrangement, and, where possible, the photographs were shown at the proper perspective for a person looking down on them in the case.

4. SIZE- Contiguous or associated items, illustrations, drawings, and backgrounds were kept to the same proportions, as far as possible, in order to link the viewing, i.e. lead on the viewer's gaze.

5. FUNCTION- Tried to place similar objects side by side for the same reason as in "4".

6. HEIGHT- Objects which illustrated the answer to a specific test question were left in place for the duration of the testing sequence so that there would be no height or eye level variation in viewing them.

7. SPACING- Objects were not allowed to intrude into the far left and top left areas of the 'Age' display panels. Those spaces were needed for large explanatory labels. Space also had to be provided for smaller labels near individual items and groups of items.

8. HOLDERS- Holders for objects were made as unobtrusive as possible. Some, such as the recesses and ledges for the burial pots, were incorporated into the design of the display panels, while others were covered in felt or paint which matched the colour of the display background.

9. FRAGMENTS- Broken fragments were either concealed or incorporated into drawings of the objects.

10. REVEALING- Important aspects of objects which could not be seen properly were illustrated beside the object. For example, the drawing of the under side of a pin, in the Late Iron Age panel.

3.3. LABELS:

A review of good labelling practice appears in Chapter 1. The criteria which governed the selection of the labels which were used in the Principal Display Treatments are amply discussed in that chapter.

The labels for the **MTs** were designed to demonstrate the "best" label format. They were concise, the text was broken-up into several short paragraphs, and paragraphs were separated by blank lines rather than by indentations. In addition, long text was broken-up into several labels, and bold headings provided a quick view of the label contents. Most of the labels used in the **BT** and **BU** displays followed basic good design practice too, however the primary labels were long and contained contiguous paragraphs separated by indentations rather than by blank lines. They also lacked bold headings and consequently took up less space than the **MT** arrangement. Although not the proposed "optimum" arrangement of text, they did allow more space for the objects and fewer distractions from them.

1. Colour Stratification of Labels within the Display

The labels in the experimental display case had different background colours which reflected their general importance and information content. The colour divisions chosen were:

- Cream –used for headings and primary information labels.
- Gold –used for supplementary information and information for particular interests.
- Green –(matching case background) short, descriptive labels placed beside items and directly relevant to them. Because they matched their background, they prevented the bitty appearance of numerous small specimen labels dotted all over the case. Distractions were also minimized by running the labels along the full length of their relevant photograph or illustration and butted up against them.

3.4. SIGN

During all of the Principal experimental Treatments, the same sign was used to direct visitors' attention towards the Introductory Case. It was positioned just to the right of the Introductory Case on the front of a large, vertical, parallelogram with a trapezoidal base. This parallelogram was approximately three feet wide and as high as the top of the Introductory Case itself. It was covered in green felt which matched the interior of that display case.

The sign proclaimed in large letters: "Introductory Case", with an arrow beneath these words which pointed towards the display case (see the photograph in Appendix B). Below the arrow was the additional legend: "Scotland Before Her Written History". This sign had a white background and black letters.

A number of previous signs had been used in Pilot Studies of visitor traffic patterns. The original *Yellow Case* had been observed both with no sign present and with one which simply stated "Introductory Case". A more unobtrusive, buff-coloured, background was also temporarily employed. However the final sign was the most successful, of those attempted in the experiment, at attracting visitors to the Introductory Case.

3.5. GOALS

> "Meaningful, challenging questions such as those which require the application of exhibit facts, concepts, and principles to new situations are far more effective in maintaining visitor attention as well as increasing learning and retention than simple questions which merely require the location of facts within the exhibit". (Lakota, 1976 p.17)

The use of overt orientation or "Goal" statements in the Principal Display Treatments was the direct result of findings by experts, such as Lakota (1976)

and Screven (1974). Such authorities maintain that question-asking labels reduce learner fatigue, that general questions should be placed just after the relevant material, and that exhibit objectives and advance organizers should be restated within the exhibit.

Such maxims are widely approved by educationalists who maintain that people learn more easily when the structure of a learning situation is clearly stated. Much emphasis in modern educational theory and practice, particularly in programmed learning techniques, has been placed upon the importance of overtly stated learning goals and orientation information so that the learner can understand what might be expected of him and what the teacher feels is important for him to gain from the instruction. Within the Principal Display Treatment Study, such orientating statements were employed in order to discover whether mobile museum audiences would use, and benefit from, such a context.

A number of *Goals*, or orientating devices and questions, were added to each of the three principal display treatments. The treatments which contained added *Goals* also gained the letter **G** added to the end of their "names". The *Goals* themselves consisted of an explanatory panel which was placed below the sign directing visitors to the Introductory Case, and two side panels placed within the display case at either end (see photograph in Appendix B.I). Both the explanatory panel and the *Goal* questions on the side panels were easily distinguished from other elements of the display by their colour, yellow.

The explanatory panel was positioned at eye-level to the right of the display case, and butted up against the Introductory Case sign. It contained an interest and attention-getting heading followed by two short paragraphs in bold, well spaced type face, which indicated how the visitors could use the display in order to discover general historical changes through time. These

paragraphs were in turn followed by a challenging sentence, in smaller type-faced italics, which pointed out the *Goal* questions on the two side panels. Below this sentence was a very simple diagram of the display case, showing its division into four "Ages", and indicating where to start viewing it in correct chronological sequence. The panel concluded with a further sentence, in comparatively small type face, which suggested why the display case was worth a long look.

The information contained in the explanatory panel appears below. However to obtain an indication of the scale of the type-face and an outline of the diagram it is necessary to refer to the photograph of the panel in Appendix B.I.

Who came here before the 'Scots'

Discover who came to Scotland or influenced life here in each 'Age' by examining the labels at the top of each time period.

By following the bands of tools, weapons, ornaments and burials across this display case you can find out how these things changed with time.

You can even quiz yourself by trying to answer the questions on the side panels at the beginning and end of the case.

(diagram)

It may be worth your while to spend some extra time studying this display case so that the rest of the objects on this floor may be put into perspective.

In general this explanatory panel tried to do three things: to interest the visitors in spending time looking at this display case, to indicate what sort of material the visitors could expect to find within the case and how it was arranged, and to suggest what the visitors could do with the information they

discovered.

The side panels within the display case were faced with green felt which matched the background of the case. When *Goals* were added the outermost edges of the panels were covered by a seven and a quarter inch wide strip of white cardboard which extended from the top to the bottom of these panels. The same bold headings appeared on both white strips, in vertical succession. Below each of the four headings was a different *Goal* question, printed on yellow cardboard which extended the width of the white strip in order to minimize any bitty appearance, and to emphasize horizontal lines drawing the viewer's eye into the case. The questions on the beginning panel of the display case challenged the visitors to look for general changes in materials while those at the end of the case called for specific answers. This structure was followed in order to reiterate information found on the explanatory panel regarding the general organization of the display case on the first panel, and follow that up with pertinent questions at the end which might encourage further contemplation.

The information which appeared on the side panels is reproduced below. The headings are indicated by bold type. Photographs of these side panels may be seen in Appendix B.I.

(Right-hand, beginning, panel)

Tools

Which group of people created the greatest change in the way-of-life here? In which "Age" did they arrive?

Weapons and defended settlements

Find out how warfare can be recognized by archaeologists.

Ornaments

Look closely at the designs on the ornaments in each "Age". In which "Age" do you feel the artwork really flourished and the greatest variety and skill was shown in decorating the surfaces of the jewellery.

Burials

What changes in religious customs do you think these Scottish burials may reflect?

(left-hand, finishing, panel)

Tools

How do we know cloth was used in the Early Iron Age?

Weapons and defended settlements

Did the Romans bring warfare to Scotland or was it here before they came?

Ornaments

When did Pictish artwork flourish?

Burials

Did you notice how people were first buried in Scotland?

CHAPTER 4

MEANS OF MEASUREMENT: OBSERVATION AND TESTING

Six separate techniques were selected to elicit information about the visitors to the museum and their reactions to particular museum situations. Broadly speaking, these fall into two major categories. The first category is composed of four "unobtrusive" methods of measurement, i.e. methods of which the subjects are normally unaware. Of these, two involve the observation of visitor characteristics, movements, and interactions; another involves timing visitor attention spans; and the fourth notes variations in the visitor's environment. The second category, on the other hand, is composed of two testing techniques. Questionnaires were administered to visitors who viewed a particular display case and a limited number of interviews were utilized to assess the validity of the questionnaires.

4.1. UNOBTRUSIVE OBSERVATION TECHNIQUES: Description and Analysis:

4.1.1. Initial Observation of Every Casual Visitor to the Gallery:

During this study, certain factors were noted at entry concerning all of the visitors to the floor who were not in organized groups. A record was kept of visitor numbers throughout the day (excluding school groups), yielding hourly, as well as daily totals, and of the size and composition of informal visitor groups. Such groups were broken down by sex and by adult versus child status. (Throughout this study the term child refers to anyone judged to be under the age of 14.) In addition, visitors' initial traffic patterns, i.e. their direction of movement from the door into the gallery, were also included in the record. (See Observation Sheets, Appendix A.)

Visitor numbers were necessary to assess the maximum possible available audience for the display case at any given time; to help gauge the effect of environmental conditions, such as weather, crowds, and quiet; to provide correlations with the length of time it took to complete the testing of each experimental display variation; and to provide comparison with the attendance figures for the entire building, i.e. both the museum and the conjoint portrait gallery. Differentiation of the sexes was necessary to demonstrate whether or not the experimental display case was attracting a significantly larger proportion of the total number of visitors of one sex (i.e. if it was male or female orientated). Group size on entering the gallery could be compared with group size in front of the display case to see if visitors were likely to retain their group formation when viewing the gallery. Group composition, in conjunction with group sex, size, and numbers, provided useful indications of some of the prominent observable characteristics of the general visiting population.

4.1.2. Observation of Visitors in the Vicinity of the Experimental Display Case:

A record was kept of the specific reactions of all visitors who looked at the experimental Introductory Display Case. Those who glanced at the case in passing were noted along with those whose interest in the case was retained for a longer period of time. A clear distinction was made between the observation of all visitors, and of those who completed the questionnaire.

Visitors were only asked to complete the questionnaire if they had looked, or glanced, at all four sections of the display case, had spent at least 45 seconds in total viewing the case, and appeared to be over the age of 14. The time limit of 45 seconds, minimum, was based on the findings of pilot studies. Visitors who viewed the display case for a shorter length of time frequently refused to attempt to complete the questionnaire, since they protested that

they had not looked at it long enough to learn anything. Indeed, the results from questionnaires which had been completed by visitors who viewed the display for less than 45 seconds indicated that they were right!

The following additional observations were only recorded for those visitors who were asked to fill in the questionnaire. For these, a record was kept of the direction from which they approached the display case, the sequence in which they viewed the four sections of the case, whether or not it was the first display case they saw, and the size of group they were with when they entered the gallery, and when they viewed the display case. In addition the number and type of interactions (i.e. verbal comments and gestures) made by visitors regarding the case was tabulated.

The observations noted above provide many interesting variables for analysis. The relative attraction value of each of the experimental display variations may be measured by comparing the numerical proportions of the available visitors who glanced at the display case, with the length of time their interest was held. Visitors' direction of approach to the display case, and their sequence of viewing it, may be correlated with their questionnaire results, in order to indicate the influence of studying the case in forward or backward chronological sequence, or in random order. In addition, the influences of entry group size, visitor interactions, and potential 'museum fatigue' (i.e. information overload and tired feet), may all be compared with the total questionnaire results to determine their effects on learning.

4.1.3. Timing of Every Visitors' Attention to the Display Case:

One method used to assess visitors' interest in the case was to measure the length of time they spent viewing the overall material in the display. 'Glances' as well as longer viewing times were recorded for every visitor in the

Gallery. When supplementary material, in the form of orientating statements, were placed on a signboard in front of the case, a separate measurement was made of the length of time visitors read that material. In order to time varying numbers of visitors who viewed the display case simultaneously, the observer used seven stopwatches arrayed in a fixed housing.

The possibility of a relationship between viewing time and test score, suggested one potential method of dispensing with the necessity of testing visitors, or of interacting with them by some means, in order to measure their learning from a particular situation. If a significant positive correlation could be found between test scores and viewing times, strong arguments could then be raised to dispense with the necessity for testing altogether (see Chapter 6, Table 6.6). By relating the length of time visitors spent reading orientating statements, to their test scores, it was possible to obtain an indication of the influence such statements had on subsequent learning and retention of information found in the case. It was also interesting to note how often visitors made use of such information when it was provided for them.

4.1.4. Daily Record of Environmental and Calendar Factors:

The calendar date of each individual segment of the experiment, and the day of the week on which it occurred, were recorded in order to assess any temporal changes in the visiting population or differences between weekend as opposed to weekday visitors. The weather each morning was also recorded, as were the general floor conditions (i.e. quiet with few visitors, crowded and noisy, or any notable conditions inside the gallery itself, such as workmen drilling). All of these external factors were recorded at the start of each day in order to gauge their effect on one or all of the following: visitor numbers, viewing times, retention of information, and general demographic characteristics.

MATERIALS:

4.2. QUESTIONNAIRE DEVELOPMENT:

4.2.1. Aims of Testing

The tests developed for this study aimed to measure specific changes in the communicating effectiveness of the experimental display case (i.e. how much the visitors learned from it) which resulted from physical changes to the display itself. For this purpose a questionnaire was developed to test the visitors' retention of, and ability to utilize, specific information found in the display case; and to learn something of their demographic characteristics and interest in Scottish archaeology. The test section of the questionnaire was created to provide the visitors with a structured format through which they could demonstrate various types of learning as well as to allow them ample freedom to express their likes and dislikes, interests and comments.

4.2.2. Selection of Questionnaire Over Other Methods of Measurement

Various principal methods of assessing learning from a display were initially considered. The most feasible choices were: an electronic 'quiz yourself board', individual interviews, and of course, questionnaires. The merits of each were then considered:

Quiz Boards– could have been set up at the end of the display to record the visitors' responses. These could have provided feedback as well as fun and a great deal of visitor participation might have been anticipated. However answering such boards often becomes an end in itself (see Screven, 1970, and

Lakota, 1976) and could be expected to attract visitors who had never even seen the display. Such boards also allow no opportunity for individual visitors to express their opinions freely. Furthermore, the possibility exists that they might be biased towards the young, who are more familiar with electronic games, in general, than are the older visitors.

Interviews– rate highly in the flexibility they offer for the positioning (and often the wording) of questions. The interviewer has an opportunity to explain the purpose of the research, to ensure that each question is understood, and to establish and maintain rapport and visitor interest. The interviewer furthermore is able to follow particular responses with probes for added information and to encourage a greater variety of spontaneous information. In addition to asking questions, the interviewer may also show pictures, objects, cards, etc. as well as clarify ambiguities in answers and classify answers as they are spoken.

The disadvantages of interviews are that they take a fairly long time to complete and often involve a number of interviewers who must be briefed, trained, organized, assisted, and usually paid. Interviewers themselves may influence the results due to personal bias and subjective interpretation of the answers. They may also influence the responses obtained due to their appearance, accent, and inflexions. Unless a great number of interviewers are utilized, the number of visitors who can be approached is necessarily limited by the length of time it takes to complete each interview. Consequently the interviewed subject group is often much smaller than the sample which can be reached through a questionnaire.

Questionnaires– which were administered by an interviewer but left to be completed by the respondent were selected as the best medium for testing learning in this experiment. Because such questionnaires could be completed at leisure by the visitors, they allowed the interviewer to continue with

observational studies at the same time. With questionnaires a reasonably large sample of visitors could be approached by one interviewer, thus dispensing with the need to hire or recruit additional staff. This combined approach has a reputation for encouraging a high response rate and ensuring accurate sampling with a minimum of interviewer bias. Additionally, the presence of an interviewer provides visitors with the benefit of initial personal contact as well as the opportunity to ask for any necessary explanations. The interviewer also is able to compare recorded information and observations about the visitor with the visitors' answers to a specifically designed questionnaire. Any attempts at cheating could also be discouraged. The interviewer need not be particularly skilled, but does need to present a consistent approach to all visitors.

4.2.3. Design and Layout Rationale of Questionnaire:

Questionnaires, as opposed to interviews, offer several distinct disadvantages. They must be much simpler, since questions cannot be amplified or immediately explained in detail. They are limited to subjects who possess adequate reading and writing skills in the chosen language and are likely to intimidate people of low intelligence or limited educational background (see Oppenheim 1972, p.33). In addition, no control can be exerted over the order and sequencing of questions answered.

In regards to the questionnaire used in this study, (see Appendix A), however, these disadvantages were minimized where possible, or regarded as insufficiently important in this context to warrant the use of a different technique. Simple wording and encouraging instructions, drawings, and a consciously informal format were used throughout. These variously attempted to interest test-wary people and those of limited education as well to provide visitors with a means of expressing nonverbally related recall. Question

interrelations were avoided as much as possible to negate the influence of the order in which questions were answered. Furthermore, preliminary pilot studies indicated that the visiting population aimed at contained a negligible proportion of adults unable to cope with completing a questionnaire of the complexity used. The further disadvantage of a rigid format was minimized by encouraging free response on a number of topics.

4.2.4. Basic Considerations in the Selection of Question Wording and Placement:

Certain general rules were followed in compiling the questionnaire. (The following suggestions for question wording and placement have been taken from Payne 1951; Kahn and Connell 1957; Selltiz 1959; Festinger and Katz 1953; and Oppenheim 1972.) Questions were kept fairly short (Oppenheim, p.56). Familiar, simple words were used wherever possible. These words were taken from lists of the 7,000 most familiar words in the English language, compiled by Thorndike and Lorge (1952). Questions were structured to require a minimal amount of writing on the visitors' part. These measures were taken in order to allow the visitors to complete the form in a short time, i.e. under 15 minutes; to make it appear less like an essay for school; and to diminish any bias in the answers resulting from visitors (both native speakers and foreigners) who were reluctant to express their thoughts due to spelling difficulties.

Since the object of a visit was presumably primarily to see the exhibits (since few other options, including toilets, were available on the floor containing the experimental display case) rather than to take a test, it did not seem fair to ask visitors to divert their attention for very long away from the objects of their visit. The opportunity existed for visitors to express lengthy opinions, but such remarks were considered optional rather than mandatory in coding the responses. A consistent attempt was followed to reflect respect for the visitor in requesting answers to the questions. The questions were not so

simple that they appeared to talk down to the visitors, neither did they make use of technical abbreviations or terms outside most native speakers' general knowledge. Oppenheim (p.65) maintains that it is necessary to continually demonstrate consideration for the visitor since "each question and response affects (the visitors') motivation and willingness to go on and give best". Carrying on from this, personalized, nonbureaucratic language was used to seek the visitors' co-operation for both the introductory statement and the statement requesting demographic data. Oppenheim also advocates the use of a minimum number of open-ended questions (p.66) but states that when they are used, no lines should be provided after them to suggest a limit to the answer.

4.2.5. Selection of Layout and Instructions:

4.2.5.1. Length-

From the outset the questionnaire was limited to two pages in length. Initial conversations with visitors suggested that a longer length would severely reduce the completion rate as well as the visitor's willingness to agree to participate in the study. However, despite this imposed constraint, extra space was obtained by suggesting that visitors continue free response questions on the back of the pages, and by utilizing a small typeface which allowed more questions to be put on each page.

4.2.5.2. Format-

The general layout of the questionnaire was aimed at producing a nonthreatening, or bureaucratic looking, quiz. Quite a few drawings were included, helping to enhance this image. Boxes and compartmentalized grid lines (see questions 1,2,3 in Appendix A) were utilized to demarcate answer choices, consolidate uneven sentence lengths, and spotlight drawings. Through such measures a generally neat, attractive, varied, interesting questionnaire of

balanced proportions was created, unlike any familiar tax form (see Oppenheim's admonition for respondent orientation a la Festinger and Katz, on page 65).

4.2.5.3. Rapport–

A variety of devices were used within the questionnaire itself to maintain some form of rapport with the visitors, encourage them to complete it, and minimize the apparent difficulty of the questions. The first question tried to immediately involve the visitors by asking them to 'tick' ( ) relatively easy statements which provided them with a twenty-five percent chance of being correct. The two questions which generally took the longest time to answer, questions "3" and "7", were placed on separate pages. Several questions were grouped together to give the illusion of single questions, i.e. questions "1", "3", and "7", thereby reducing the apparent number of questions being asked. These "multi-questions" were also spaced out within the questionnaire, as were the interestingly illustrated questions, i.e. questions "2", "3", and "6". Additionally, respondents were encouraged to continue the questionnaire on the second page because an attractive-looking illustrated question was positioned at the top of that page. Visitors were issued a personal challenge as incentive for them to undertake the last, and rather difficult, question. (Judging form comparisons with other approaches used in pilot studies, these measures significantly enhanced the completion rate of the questionnaire, i.e. they were effective.)

4.2.5.4. Instructions–

The instructions to the visitor were aimed at clarifying the questions, maintaining rapport, and offering encouragement. The questionnaire was headed by a direct personal appeal which contained ample encouragement and a brief explanation of the purpose of the experiment. This approach to the respondents specifically avoided mention of confidentiality and sponsorship,

since the overriding constraint was to present an informal, non-hostile, even enjoyable (!), document. The two sentences in the introduction were kept short and simple and they were separated by indentation, and centred within a five inch space in the middle of the page to facilitate skimming (see Williams, 1960), and to draw attention to them since introductions are often overlooked (own experience).

Single instructions for 'ticking', 'numbering', and 'linking' were used to cover the "multi-questions" mentioned previously. This format ensured that a good deal of information could be elicited in the available space, and encouraged the illusion that the respondents were only answering seven questions. A wide variety of instructions were used, to 'tick', 'fill in the blank', 'number in sequence', 'circle', describe 'interests' in the display case, and 'draw a line between' (i.e. link) in order to encourage respondents to adopt a fresh approach to each set of questions (as opposed to the centralizing tendencies encouraged by 'ticking' down a form). Instructions which asked questions or 'challenged' the visitor were interposed between those soliciting cooperation. Descriptive statements were occasionally used (questions "2" and "3") to provide information about illustrations before presenting a problem to be solved. Oppenheim's recommendation to use a further introductory phrase (p.58) to explain the necessity for answering demographic questions was followed.

4.2.6. Recording Descriptive Data on the Questionnaire:

The upper corners of the questionnaire were reserved for the interviewer to record pre-coded information which identified each individual questionnaire and the date on which it was administered to a subject. The upper right-hand corner was used to designate each questionnaire utilized during a particular treatment condition. Following the abbreviation "**no.**", the number of the questionnaire's placement in the ordering from one to two hundred was written

in. This number was used to correlate the answers to each questionnaire with the observational material recorded about the subject on the data sheets. Two hundred questionnaires were utilized for each principal test variation, hence the cut off point at 200 in the questionnaire numbers. Below the abbreviation for 'number' is the abbreviation for the experimental treatment variation, "**var.**". Beside it would be written the code for the particular treatment condition to which the visitor was exposed, for example, "**BU**".

The upper left-hand corner was used to denote the day of the week, the month, and the day on which that questionnaire was given to a subject. On the top line, the interviewer had the choice of circling one of the following: "**S S W**", to indicate that testing took place on a Saturday, Sunday, or Weekday, respectively. Specific weekdays, when they occurred, were written in conventional abbreviations to the right of this line. The first letter of the relevant month could be circled in the line below and a blank was provided for insertion of the day of the month. This second line was written "**M A M J___78**", with the letters standing for the months during which the experiment was expected to run, i.e. March, April, May, and June respectively. Since testing actually continued into July, the final "J" was also used to indicate this additional month by underlining, rather than circling, it. Decipherable abbreviations were utilized for this descriptive information so that they could conceivably be puzzled out by the subjects if they so wished, thereby hopefully avoiding the look of the incomprehensible codes which often appear on large computer coded questionnaires.

4.2.7. Open-ended Questions and Concluding Statements:

Both pages of the questionnaire were terminated by an open-ended, personal opinion question, directed at the respondents. These questions were kept rather broad in order to encourage a wide range and quantity of

responses. They were specifically aimed at highlighting the most memorable aspects of the display, clarifying the type of 'visitor-referenced' learning, and illuminating the areas of display which visitors themselves felt helped them to extract what they wanted from the display case. In short, these questions attempted to add depth, 'color' (via expression), and breadth to the more quantifiable fixed format responses.

The first question which allows visitors real freedom of expression is number "5". In it, visitors were asked:

"**What did you find in this case which particularly interested you-**"

This question was terminated by a hyphen (-) rather than by a conventional question-mark (?). The choice was a subjective one, intended to imply that the question should be added too, or continued, since it displayed no recognizable end point. Below this question, over two inches of space was left free for the individual's reply.

The other free-format question, or more precisely, question group, was hand printed and left without a designating number or letter, thereby differentiating it from all other questions in this study. Such informality was employed, again subjectively, to stress human involvement in the question and interest in the reply, thereby hoping to avoid a terse, form-filling, type of response. Studies of visitor behaviour in museums, such as the Peabody Museum study (Kearns, 1940) suggest that marked changes in visitor response may be attributed to personal direction. The personalized layout of this question, together with its heading, "**Comments**-", was also used to suggest the optional nature of the potentially time-consuming response, and obviated the recommended use of preceding statements such as, "Please disregard any questions you do not wish to answer."

This question was also carefully positioned within the questionnaire to emphasize the quantity and nature of the hoped for response. Following, as it did, a 'rest-period' of comparatively easy demographic questions, it was hoped that respondents would be more willing to answer a detailed request for information. Furthermore, it was placed at the end of a rather difficult to answer questionnaire. This placement was intended to stimulate visitors to criticize the display after having been forced to give due consideration to their own learning, or lack of learning, from the case.

Criticism of the display was further encouraged by effectively breaking the question into two parts, using a filter question to set the stage for the more important question to follow. Thus the complete, handwritten, question group read:

"**Comments- What did you think of the arrangement of the display in this case? How would you like to see it improved? -(use back if needed)**"

Criticism was encouraged as a means of pointing out aspects of the display which the visitors judged to be important. A specific directive to the respondents, to use the back page if necessary, was added at the end of the question group to encourage a full response.

Two statements in the bottom right-hand corner of the second page, both handwritten, concluded the questionnaire. "**Please return questionnaire to box provided**" was the final instruction to the visitors. It was handprinted in the space below, and to the right of, the question group discussed above, in order to obliquely remind visitors of the confidentiality of their responses (thereby encouraging criticism and minimizing ill-feeling which might result from the belief that many of their answers were incorrect) as well as to request visitors to return, rather than retain, their questionnaires. At the very bottom of the page was the courteous, script-written message: "**Thankyou very much for your**

help". Again, it was handwritten as a way of personalizing the statement. Script was used to distinguish it from questions and instruction statements.

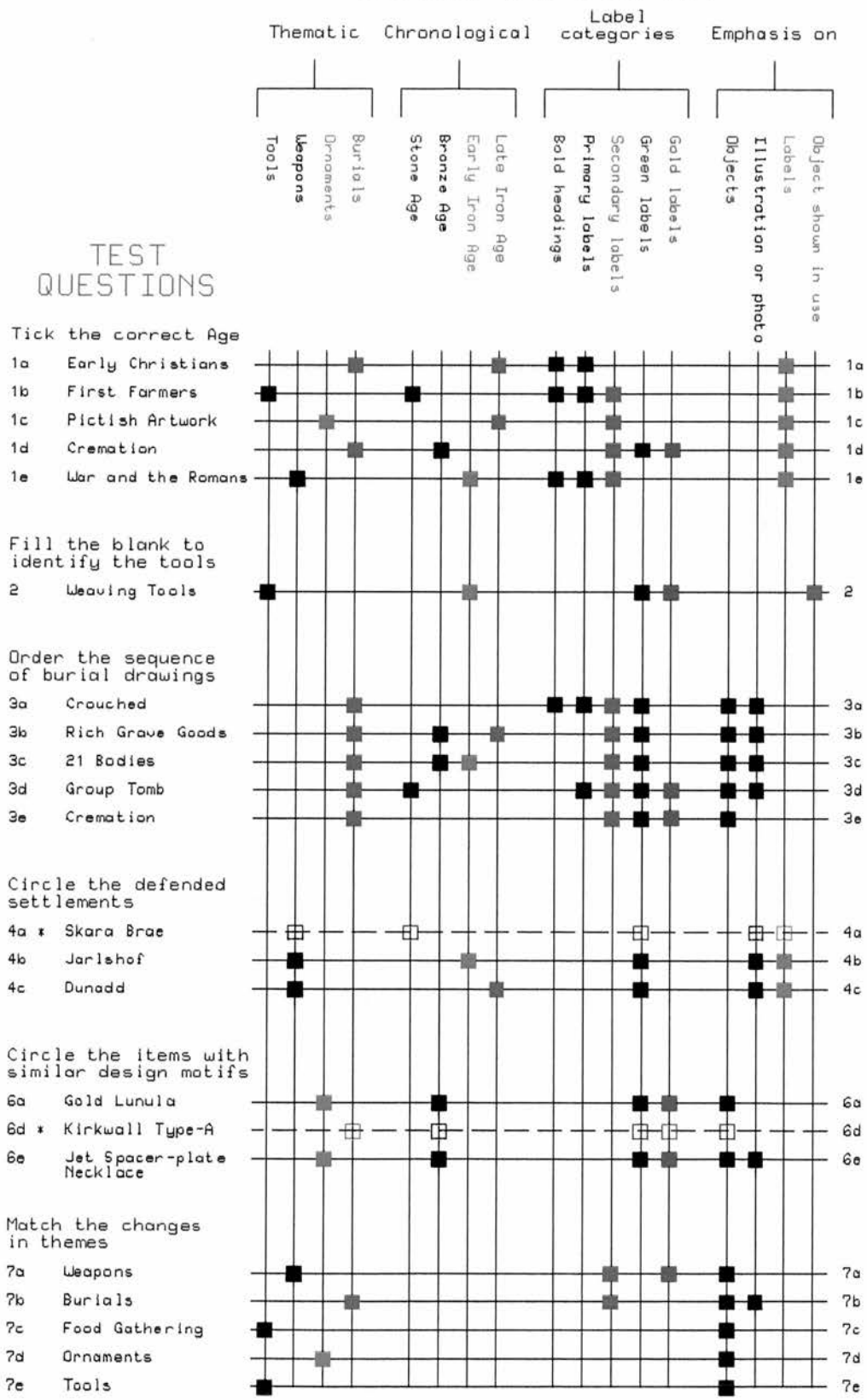
4.2.8. CONTENT QUESTIONS:

The test component of the questionnaire for this study was designed as a short, factual inquiry into the general areas, range, and quantity of information which visitors could recall from the experimental display case (see Appendix A). An attempt was made to assess demonstrable learning levels (see subsequent discussion of Gagne's Learning Stages) yet at the same time maintain an even distribution of questions between the major divisions of the exhibit case. The design of the exhibit case emphasized four chronological "Ages" and four material "Themes". Through carefully thought-out multi-dimensional questioning, these eight divisions, which also corresponded to eight horizontal and vertical areas of the case, were each represented by at least three questions. For example, one question concerning the "Theme", Ornaments, asked for information only found in the Bronze "Age". Thus an "Age" as well as a "Theme" were both represented by one question. The Coloured Chart on the next page summarizes each of the test content questions on the left-hand side and indicates where the principal sources for the correct answers may be found within the exhibit case.

The test was also designed to assess knowledge gained from viewing objects as well as reading labels (see colour chart, "Emphasis On" categories). Therefore an attempt was made to balance the principal sources for the answers between those which one was likely to discover from labels and those learned from observing the objects, illustrations, and photographs.

Overall the test content questions were designed to be interesting, varied, non-threatening, encouraging, and fun for the visitor. This test was also

DISTRIBUTION OF ANSWERS WITHIN THE DISPLAY



*Note: 4a is an incorrect response (it was defended).
6d is a dummy response (not shown in the display).

formatted to cater to different types of visitor reactions. Most of the questions required minimal written response and therefore could be completed by those who were in a hurry or who were not particularly interested in contributing their responses to this study. However those who had more time or were more enthusiastic about contributing, were provided with ample space and encouragement for additional open-ended responses.

The items were arranged in a nonthreatening format which allowed a multiple choice type of response to all but one of the test items. That test item asked for a "fill-in-the-blank" response. The multiple choice questions provided a variety of activities to keep the visitors' brains working. They could "tick" ( ) the correct response, circle their preference, number contiguous items in chronological order, and match statements with their appropriate headings.

Obviously, multiple choice questions, when compared with open-ended or fill-in-the-blank questions, increase the probability for correct guesswork. However, they were chosen: for the advantages they offered in easing the subjects into a testing situation without intimidating them; because they required less written and time effort; because they produced no shameful awareness of poor spelling habits; and because they did not blatantly show up the visitors' lack of knowledge, thereby leaving them with a sense of dissatisfaction with themselves. Multiple choice questions were also easier to code for analysis.

4.2.8.1. Discussion of Gagne's Eight Types of Learning (1973)

In addition to covering different areas of interest within the display case, the test also tried to assess, in general terms, the existence and extent of different types of demonstrable learning. The levels of learning looked for in the test may be related to several of Gagne's "Eight Types of Learning". He distinguished between (1) signal learning, (2) stimulus-response learning, (3)

chaining, (4) verbal association, (5) discrimination learning, (6) concept learning, (7) rule learning, and (8) problem solving. For a full discussion of the conditions distinguishing between these categories, see Gagne, The Conditions of Learning, 1973. The "Types" looked for in this test are briefly described below in operational terms relevant to the museum:

- (2) **stimulus-response**, which emphasizes the process of discrimination (Gagne 1973, p.38). For example- The visitor learns to call one group of items tools.

- (4) **verbal association**, verbal linking of new words with old, known words, through personal mnemonic relationships. Example- The visitor is able to correctly designate a crescent moon shaped object as a lunula because the name is similar to the Latin word for moon, luna, with which he is already familiar. (see Question 6)

- (5) **multiple discrimination**, the ability to call objects or groups of objects by their individual names (Gagne 1973, p.50). Example- The visitor, given a short list of settlements, will be able to correctly identify the defended sites. (see Question 4)

- (6) **concept learning**, internalized representation of a set of concrete or defined items. Example- The visitor is able to select the odd item, such as a weapon, from a list otherwise composed entirely of tools.

- (7) **rule learning**, a chain of, or relation between, two or more concepts (Gagne, p. 56-7), or the acquisition of the "idea" or principal contained in a designated proposition. Example- The visitor, given the proposition that there is no evidence that warfare existed in the Stone Age, is able to correctly indicate that no weapons or defended sites are known from that period.

- (8) **problem solving**, with a goal in mind, the learner is able to combine relevant, previously learned rules to generate a new rule. Example- Having recalled representative examples of burial types found in each "Age", the visitor is able to put these together to gain an overall understanding of changes in burial customs through the Ages. (see Question 7)

In all of the above descriptions, it should be noted that Gagne's basic argument is that "...different varieties of learning can be distinguished by the conditions required to bring them about." (p.56) Although a number of the "Types" above are not directly, or solely, related to a particular test question,

they describe some of the many and complex processes the mind must go through inorder to arrive at an acceptable answer. These learning classifications are also directly observable through the visitors' *Written Comments* as with the example cited for "rule learning".

4.2.9. DEMOGRAPHIC SECTION:

4.2.9.1. Layout

The nine demographic data questions are grouped together on the lower half of the second page of the questionnaire. They are clearly separated from the standard content questions, "1"-"7" above, by a dashed line which runs across the width of the page. These demographic questions are further distinguished by labelling them with the letters "A" to "I" rather than by continuing the numbered headings used above. Such alphabetical headings are useful as a means of disguising the actual quantity of questions asked.

4.2.9.2. Location

Several considerations were sufficient to dictate that demographic questions be positioned at the end of the questionnaire. By leaving such 'non-content-based' questions until the last, it was hoped to maximize the opportunity for visitors to immediately recall information acquired from the display case and to minimize possible obscuring effects. (Attempts were especially made to minimize distractions since measurable detailed learning from a museum display has rarely been satisfactorily demonstrated with adults.) Furthermore, the possibility that some visitors might find the demographic questions offensive suggested that they might consequently refuse to complete the remainder of the questionnaire if the demographic questions were placed at the beginning. In addition, the layout which was used seems more logical than vice versa. The emphasis is thus placed on the more important questions first, thereby establishing the 'worth' of the questionnaire

initially and substantiating the appeal given at the head of the demographic section requesting visitors to supply demographic details.

4.2.9.3. Introductory Statement

A rationale for all of these questions which seek information about age, sex, education, occupation, etc. is presented to visitors at the head of the demographic section. This was hoped to act as an inducement for them to answer the questions truthfully and thoroughly. The rationale reads:

```
Please answer some questions about yourself so that we
can find out if a representative sample of visitors are answering this
questionnaire:
```

In addition to seeking visitors' co-operation, this explanation also serves to emphasize the point of view that the respondent is a thinking person, who deserves an explanation for his effort, rather than an ignorant guinea pig who simply does as he is told.

4.2.9.4. Space Constraints

The constraints imposed upon the entire questionnaire by the shortage of space for a wide range of questions and the desire for minimal answering time were especially noticeable in this section where the questionnaire layout required that sufficient free space be left for visitor comments below. The epitome of abbreviated instructions occurs just beneath the explanatory introduction quoted above and occupies space between sentences which would otherwise be left blank. This hieroglyphic, "( )", was inserted to indicate that answers should be "ticked" within the appropriate brackets. It also transcends the international differences in terminology between the British term 'ticking' and the American term 'checking'.

With sparing use of space, four of the nine demographic questions take up less than a single line of type. Three occupy two lines each, and only two

questions take up three lines. Occasionally, two questions are potentially asked in one, as in questions "A" and "I", where the answer to one question may lead to another question being asked. Likewise, the resultant eleven possible answers to the questions are equally terse.

Five of these answers require "ticking" (questions A, B, C, E, I) and three are essentially "fill-in-the-blank" (questions A, G, H). However where more detail is requested (questions D, F, I) the questions simply end in a hyphen followed by free space. All questions are single-spaced within.

4.2.9.5. Statement of and Justification for Each Demographic Question:

The first demographic question, "A", asks:

```
Is this the first time you have visited this museum?
( )Yes.  ( )No- then please estimate the number of times you've visited
                this museum in the last 12 months _____________________
```

This question was constructed to give an indication of the number of repeat visits to the museum. From such a question, an assessment may be made of the proportion of repeat visitors which may be expected. The likelihood of return visits presents strong implications for both the potential for second viewings of the displays, and, when combined with the range and mean for the viewing time, for the degree to which secondary labels are likely to be read. A definite time period was emphasized within the second part of the question in order to minimize ambiguities when asking for recurrent forms of behaviour (see Moser, p.57, and Appendix A). Punctuation, in this instance a full stop and a hyphen, were used, respectively, to indicate when one answer was sufficient, or when it lead into another question. (This same device was similarly employed in question "I".) Spacing was effectively used to mark the segregation of the supplementary question, which followed after the "No" response, from any proximity to the "Yes" response.

Questions "B" and "C" sought to uncover the respondents' subjective impressions of their own interest in, and knowledge of, Scottish archaeology. These questions appeared as follows, with "B" before "C":

How interested are you in the early history and archaeology of Scotland?
( )very interested, ( )moderately, ( )slightly, ( )not interested

Would you describe your general knowledge of Scottish (or British) archaeology as- ( )very good, ( )good, ( )fair, ( )poor, ( )non-existent

Certain considerations prompted the inclusion of these questions. Interest factors play a large part in most current motivational theories of learning. Thus it was felt that correlations made between avowed interest and viewing time might provide some analogy with broad test scores, thereby suggesting that sources other than a test could be used to indicate broad levels of learning from a museum case. An indication of the general knowledge which visitors brought to the display case (as opposed to knowledge gained from viewing the case itself) was felt to be useful in deciding upon the amount of detail, and the number of tangential items, to include in future displays. As a result, if visitors were shown to generally possess a firm grounding in the basic framework of Scottish archaeology, then future displays would aim to expand and elucidate this. However, if most visitors were found to come into the museum with little previous basic knowledge, then the initial displays might be expected to stress and reiterate the basic concepts.

Other factors influenced the layout and choice of wording within these questions. Question "B" offered a choice of four answers, whereas "C" offered five, since it was felt that interest was somewhat easier to classify than knowledge. Space was particularly conserved in Question "B" by deleting the word "interested" from the two choices in the middle of the answer group, without destroying their meaning. Superlatives, such as "excellent" were left out of the answer choices since it seemed likely that modesty would inhibit

their selection (see Appendix A) while facetiousness might enhance their choice.

A precise indication of visitors' places of origin was sought in question "D" (below) in order to indicate the major catchment areas for the museum, i.e.:

```
Where are you from?
Edinburgh residents- please give district or street-
Others- please give country and town or state-
```

The desire to provide an estimate of the proportion of visitors who lived within certain radii of the museum prompted the request for district or street names within Edinburgh. In phrasing the leading query, "Where are you from?", 'from' was the operative word, used to emphasize potential lack of knowledge, language, and terminology difficulties suggested by visitors' continued identification with their point of origin, despite the possibility of a different current address. Thus, knowledge of visitor-defined "home" was given some preference over knowledge of travelling distances to the museum (which could have been solicited through use of an alternative leading question, such as, "Where do you live now?"). Consequently, in the space allocated for one residence question, preference was given to the question wording which was more likely to imply possible learning difficulties rather than to supply information for transportation studies.

Questions "E" and "G" ask for straight forward information about sex and age, as seen below:

```
E.   Are you- ( )Male.  ( )Female.

G.   What is your age now-_____________
```

Question "G" was particularly phrased so as to include the word "age" as a lead-in to the more awkward question "H" (see below):

H. At what age did you stop being a full-time student-_____

The age when schooling was completed was selected as a fairly inoffensive indication of the visitors' general educational attainments. The use of such ages transcends the difficulties of comparing specific degree qualifications, both within and between countries. It exempts visitors from the embarrassment of being asked to list degree qualifications which might not exist. Specific educational achievements were not deemed to be of exceptional importance, beyond their potential implications for learning in regards to the duration and depth of past full-time study. Although the age when schooling ended may be used to imply the former, it has the added advantage of allowing an easy calculation of the "resentness" of full-time education, with its consequent implications for current learning. The phrase, "full-time student", was especially employed to distinguish students on actual courses from people who studied on their own or who were only partially involved in a student atmosphere.

Question "F" asked:

What is your occupation and what exactly do you do in it-

Provided the additional information is included, responses may be more precisely categorized than is often possible with single-word occupation replys. For example, if the reply is succinctly stated as "Engineer", the visitor may be occupied as anything from a train driver to a scientist, planner, or electrician.

Question "I" asked:

Did you know you were going to be asked to fill-out this questionnaire?
( )No. ( )Yes- then how did you know-

Finally, question "I", above, attempted to ferret out any prior knowledge that the questionnaire would be administered, because of the attendant implication that

the case would be more closely studied, thereby biasing the results. However, it was discovered that the rare visitor who did realize that people were being questioned about the case, more often than not, simply moved away to avoid being tested rather than begin an in-depth examination of the case.

4.2.9.6. Rationale

Demographic questions provide information to the designers and curators of the museum which is useful in helping them to plan the displays. A knowledge of the audience, its range of capabilities, previous subject knowledge, and likelihood of returning, are important in accurately pitching the labels and material. Furthermore, it is important to understand "who" the audience is before jumping to judgements about what the audience should be.

The questions asked provide basic information about certain museum visitors, i.e., their frequency of visits, residence, sex, age, education level, occupation, and knowledge of and interest in, Scottish archaeology. Such information can be used to demonstrate whether comparable samples of visitors were tested across the experimental treatment variations. Specific comparisons between demographic variables and question response may also be made. For example, comparisons can be made between avowed interest, knowledge, and education levels, and, the full completion of the questionnaire, the score, the attention span, etc. In addition, the results of the demographic questions make it possible to divide visitors into categories for separate analysis and to compare their overall characteristics with other museum surveys.

Despite the interesting aspects of the results of the demographic questions, it is important to keep in mind that, apart from the control groups, these questions were intended to be applied only to a selected segment of the museums' population, i.e. those who viewed one display case. Since the

results would not necessarily be conclusive for the museum as a whole, a decision was taken to limit the demographic questions to those which were particularly suitable for comparison with the results of the 'content' questions or for influencing the museums' plans for future displays. Occupation, interests, education, etc.– a barrage of questions would be necessary to fully cover each of these topics. However, within this questionnaire, each of these topics is allocated only to one question. Obviously such questions will not, for the most part, elicit a great depth of information. But, even so, their influence may be extended beyond the 'learning orientation' of this study to indicate whether the range of audience characteristics warrants a renewed evaluation of the museum's planning and display policy as a whole.

4.2.10. VALIDITY:

Several methods of judging the validity of the questionnaire were incorporated into the format of both the questions and the overall experimental design. Many of the methods were suggested by Oppenheim (pp. 69–72). Proof of this factor is essential in establishing the worth of the questionnaire. If a questionnaire is valid, then the questions within it are known to measure the right things (i.e., they measure what they purport to measure).

One exhibit variation was tested twice within the experiment as a cross measure of the validity of the questionnaire. This variation was tested at both the beginning and the end of the study, and was designated as **MT** and **RMT** respectively. Thus, allowing for possible differences in the visiting population occurring during that timespan, analysis of the answers of comparable segments of the respondents should yield similar results.

Five principal methods of measuring the validity of the questionnaire were employed in this study. Many of the methods were suggested by Oppenheim.

These were: interviews, observation, pre-test versus post-test answers, and the use of 'pre-case' and 'at case' control groups. The separate criteria thus established each offered an independent measure of essentially the same variables.

In the first method, scattered interviews were conducted with visitors who had seen the display and who had completed the questionnaire. These interviews covered the same material as that found in the questionnaire itself. They were useful in establishing the validity of the questionnaire by demonstrating whether or not visitors actually understood the display content behind their correct responses to the questionnaire and were unsure of the material behind their incorrect replys.

The second measure of validity occurred in the cross-check used between observations and answers. Overheard visitor comments which specifically pertained to questions in the survey were noted and were later compared with that visitors' responses on the questionnaire, in order to establish whether the relevant questions were correctly answered.

Comparison of the results of the pre-test versus the post-test answers provided the third measure of the questionnaire's validity in measuring visitor's learning from a display case. At designated times, all of the visitors who entered the gallery were approached to participate in a special experiment. They were asked to answer the questionnaire before they saw the display case, view the case for as long as they wished, then complete questionnaire again. By comparing the same person's responses on both tests, conclusions may be drawn regarding the amount of learning which resulted from viewing the experimental display.

The fourth and fifth measures of validity were provided by two differently directed control groups. The first group was composed of visitors to the

Scottish archaeology section of the museum who were asked to complete the questionnaire without seeing the display case. This group included the visitors who answered the pre-test questionnaires mentioned in the preceding paragraph. The answers provided by this "pre-case" group were used to establish a base limit of the proportionate correct responses for each question asked. Thus this base line may be seen as an estimate of the amount of information which visitors were likely to bring in with them into the museum, or be able to guess correctly, concerning the topics covered. A particular question was deemed to be invalid, and was consequentially discarded, if more than a third of such 'uninitiated' subjects were able to answer it correctly.

The second control group consisted of visitors to the gallery who were asked to complete the questionnaire while they were standing in front of the test display case. They were the only subjects in this study who were allowed to study the case while they were answering the questionnaire. This group was used to establish whether it was indeed possible to answer the questions correctly based on information found in the display. Ideally, all subjects in the first group would receive a score of "0", while all those in the second would obtain the maximum possible score. Reality, although significant, was a step behind this ideal (see Appendix G.I).

On a small scale, a split ballot pilot study was used to measure ordinal biases. Non-existent items were also included in the possible responses to the content questions.

CHAPTER 5

CONDITIONS UNDER WHICH THE TEST WAS ADMINISTERED AND GENERAL DEMOGRAPHIC CHARACTERISTICS OF THE VISITORS TESTED

The National Museum of Antiquities of Scotland attracts people from many different countries, educational backgrounds, and occupations. These visitors arrive with a varied knowledge of, and interest in, Scottish archaeology and antiquities. They come on different days of the week, at different times of the day, in different months, and under different weather conditions. When they enter the museum, they may find themselves in a noisy environment full of boisterous school children or in a quiet haven with few other people about. This chapter sets out to determine whether there were any significant differences in the environments under which the visitors who were tested looked at the six main display treatments (previously referred to as **MT, MTG, BT, BTG, BU, BUG**) or in the demographic characteristics and knowledge and interest levels of the tested visitors themselves.

The first part of the chapter deals with the observed test conditions, i.e. some of the external conditions under which the visitor sample looked at the display treatments and were tested on the information which they retained. These observed test conditions comprised: the day of the week that testing took place, the number of visitors tested on each day, the number of potential test subjects available, the number of days and hours it took to test each treatment, the weather and floor (noise and crowd) conditions present during testing, and the size of group the testee was in when s/he entered the First Floor Gallery. These test conditions were primarily observed in order to ascertain whether they differed significantly between treatment groups. If major differences did occur, it is important to attempt to assess whether or not these differences are likely to have affected the evaluation measurement

results, which are discussed in Chapter 6.

The second part of this chapter deals with the testees' responses to demographic and interest questions. This information was volunteered by the testees after they had completed the test and the questions were kept to a neutral minimum so that most of the subjects would answer them. Because of the general nature of these questions, the answers to status related topics, such as occupation and educational level, were often vague and difficult to classify with any degree of certainty. Consequently these responses must be taken as merely a general indication of the relevant characteristics of the testee sample.

5.1. TEST CONDITIONS:

5.1.1. Day of the Week Testing Took Place and Sample Size (number of visitors) Tested on each Day:

The decision was taken to test the maximum number of visitors on each day (i.e. every person of fourteen years of age or older, who looked at the display for at least 45 seconds) in order to complete the entire testing procedure as soon as possible, rather than to test a predetermined number of people on each day of the week (as in the ROM study among others). This time restriction was based on two factors. One was the need to finish the study within the projected start and completion times of concurrent construction work within the museum. A further reason was the a priori assumption that the demographic differences between the testees were likely to increase with the advent of the summer tourist season, and to be of more consequence than the differences between the samples tested on different days of the week. The second assumption certainly appears to be inaccurate (see the discussion of the chi square results following table 5.1, which shows a

difference between weekend versus weekday visitors.) There were also variations in such expected seasonal differences as residence, but these differences did not follow a set or predictable pattern which would suggest a gradual increase in summer tourists. In fact, the first and last display treatments tested had similar subject residence populations, which were different to those of the other treatments (see the full discussion of residence changes following table 5.9). However the temporal differences in visitor characteristics appeared to have negligible significance in their influence upon the results (see the discussions of the individual visitor characteristics in this chapter).

Because it was necessary to test the sample visitors quickly, no effort was made for consistency in deciding on which days of the week testing would take place or how large a sample would be tested on any given day. The number of visitors tested on each day was arbitrarily determined by the number who voluntarily looked at the relevant display treatment for forty-five seconds or longer, and who agreed to take the test. However it should be recognized that the latter was not a valid consideration, since only five people out of approximately 1000 visitors who were asked to complete the test, refused to do so. Therefore the real criterion for determining how many people were tested on each day was the actual number of visitors who looked at the treatment for at least the minimum amount of time.

Most display treatments were tested over six days of the week, i.e. Monday, Tuesday, etc. Some days were repeated (i.e. two Wednesdays or two Sundays and two Mondays) and different days were omitted during the testing of the different display treatments. The exceptions were **MT**, which was tested over five days of the week, and **BTG**, which was tested over all seven. The observed evaluation of a particular display treatment ended after the administration of its one-hundredth test. It took from six to twelve days to

collect a sample of 100 tested subjects for each display treatment.

Within an individual display treatment, the greatest number of people were tested on a Saturday (22.7%), and the least number on a Thursday (9%). The rest ranged from 11.7% on a Sunday afternoon up to 15.8% on a Friday. 69% of the visitor sample were tested for **MT** on a Saturday and a Monday; 68% for **MTG** on a Saturday, Tuesday, and Friday; 69% for **BT** on a Monday, Friday and Wednesday; 61% for **BTG** on Tuesday, Thursday, and Wednesday; 75% for **BU** on Friday, Saturday, and Wednesday; and 59% for **BUG** on Sunday, Monday, and Saturday. All of the treatments were tested on a Saturday and a Thursday. All but **BT** were tested on a Sunday, all but **MTG** were tested on a Monday, all but **MT** and **BU** were tested on a Tuesday, all but **MT** were tested on a Wednesday, and all but **BUG** were tested on a Friday. The table from which the detailed information in the preceding paragraph was based (see Appendix **H**) was not deemed to be sufficiently important to reproduce here. Instead, the days of the week were amalgamated into two categories, weekend– Saturday and Sunday, and weekday, for inclusion in the subsequent table.

TABLE 5.1- Number of weekend and weekday testees, number of hours and days for testing, total number in sample pool:

DISPLAY TREATMT	NO.OF DAYS	NO. IN WEEKEND SAMPLE	NO. IN WEEKDAY SAMPLE	NO. OF HOURS FOR TESTING	TOTAL NO. OF VISITORS PRESENT (Sample Pool)
MT	8	66[HI]	34[LO]	39	434
MTG	10	31	69	50	505[HI]
BT	10	10[LO]	90[HI]	45	447
BTG	12[HI]	19	81	55[HI]	500
BU	7	32	68	31	456
BUG	6[LO]	48	52	18[LO]	371[LO]
TOTAL	53	206(34%)	394(66%)	238	2713
MEAN	9	34	66	40	452
RANGE	6	56	56	37	134

CHI SQUARE FOR WEEKEND VERSUS WEEKDAY ONLY= 90.18767
WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE= 0.0000

The days of the week that testing took place were grouped into weekend (Saturday and Sunday) and weekday (Monday to Friday) categories. When looked at in this way, the chi square test (see table 5.1) shows a significant difference between the treatments. **MT** contained the highest proportion of subjects in the weekend sample (66 compared with the mean of 34) while **BT** and **BTG** had the highest proportion in the weekday group (90 and 81, respectively, compared with the mean of 66).

The total number of casual visitors, who were on the floor when a particular treatment was being displayed, made up the potential number of subjects who could have been tested, if they had looked at the display treatment for 45 seconds or longer, i.e. the sample pool. It can be seen from Table 5.1 that the number of days and hours taken to collect the requisite sample for each display treatment are directly related. Furthermore, the total

number of visitors present in the sample pools during those hours and days, are roughly in the same proportions between groups. Therefore it is generally true that the more hours that were taken to collect the sample, the more days that were needed, and the more visitors that were present in the total sample pools. Expressed in operational terms then, a display treatment which took a relatively small number of hours to attain its sample quota, also took a small number of days, during which time a small total number of visitors were present. An explanation for this apparent anomaly, that the fewer people there are in the environment, the quicker the sample quota is gathered, may be uncovered by looking more closely at the sizes of the daily sample pools.

The greatest number of visitors who entered the gallery on any single day when the treatments were being tested was 122. Using this figure as the upward boundary for the daily sample pool, the number of visitors per day may be divided into three groups of equal intervals, in order to facilitate a general overview of the situation. This grouping also makes it easier to assess whether there were any major differences in the sizes of the daily sample pools between the treatment variations. In the following table the 600 visitors who were tested are grouped according to the sizes of the daily sample pools from which they were drawn.

TABLE 5.2- Number of testees for each treatment drawn from small, medium, and large daily sample pools:

DISPLAY TREATMENTS	4-43 VIS. in SMALL D. SAMPLE POOL	46-82 VIS. in MEDIUM D. SAMPLE POOL	87-122 VIS. in LARGE D. SAMPLE POOL
MT	25	32^{LO}	43
MTG	38^{HI}	38	24
BT	34	66	-
BTG	32	68	-
BU	15	35	50^{HI}
BUG	3^{LO}	69^{HI}	28
TOTAL	147	308	145
% (and *MEAN*)	25% (25)	51% (51)	24% (24)
RANGE	35	37	50

ALL ROW TOTALS= 100 (16.7%)

CHI SQUARE = 158.83737 WITH 10 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

Roughly equal groups of daily visitors were created, with 87 to 122 in the large group, 46 to 82 in the medium-sized group, and 43 and under in the small group. However it is important to understand that this small group sometimes merely reflected the number of visitors observed when only a few subjects were needed to complete a treatment sample, or when a full day of observation and testing was prevented by outside circumstances. Both **BT** and **BTG** had a maximum of 82 visitors to the floor on any one day. Consequently their samples contained no visitors in the large group. However 50% of the subjects tested in **BU**, and 43% of those tested in **MT**, fell into the large group category. In the small group (under 43 people) both **BU**, with 15 subjects, and **BUG**, with only 3, were well under the 32 subject mean of the other four cases.

One might expect to find a positive relationship between the total number

of visitors observed on each day and the number of days taken to collect the requisite sample. This contention appears to be directly borne out in this study if one looks at those treatments which had the lowest numbers of subjects tested from the small daily sample pool group, i.e. **BUG, BU**, and **MT**, since these took the least number of days to complete, rather than by looking at the ones which had the most subjects in the large group (see Tables 5.1 and 5.2). Therefore, it would appear that the availability of medium to large numbers of people in the daily sample pool, within these limits, was an important factor in determining the length of time taken to collect a sample. The number of people present can not only increase the number of potentially interested subjects, but can also help to mask the sound of instructions being given in a quiet gallery, which might deter some people from approaching the test area.

5.1.2. Weather Conditions:

TABLE 5.3- Number of visitors in each treatment group tested under different weather conditions:

DISPLAY TREATMENTS	% OF SUN	% OF CLOUDS	% OF RAIN	ROW TOTAL
MT	80^{HI}	17	3	100
MTG	21^{LO}	38	41^{HI}	100
BT	43	57^{HI}	-	100
BTG	87	13^{LO}	-	100
BU	32	47	21	100
BUG	69	31	-	100
TOTAL	332	203	65	600
% AND *MEAN*	(55%)	(34%)	(11%)	(100%)
RANGE	59	44	41	

CHI SQUARE= 241.03394 WITH 10 DEGREES OF FREEDOM, SIGNIFICANCE=0.0000

The weather conditions, which were observed at the beginning of the day, (generally noted at approximately 10 a.m.) varied a great deal between the tested display treatments. (And the .0 significance level suggests that these variations were significant.) However, only two of the treatments contained a major proportion of visitors who were tested on rainy days. These treatments were **MTG**, with 41% tested under inclement weather conditions, and **BU** with 21%. There were three treatments which were tested under predominantly clear conditions: **MT**, 80% sunny; **BTG**, 87% sunny; and **BUG** with 69% of the subjects tested on sunny days. The visitors to **BT** were tested under predominantly cloudy conditions (57%), while **MTG** and **BU** were tested under mainly cloudy or wet conditions: **MTG**-41% rain, 38% cloudy: **BU**- 47% cloudy, 21% rain.

5.1.3. Floor Conditions:

TABLE 5.4– Number of testees per treatment who were tested under different floor conditions:

DISPLAY TREATMENTS	% OF SAMPLE IN QUIET CONDITIONS	% OF SAMPLE IN NOISY CONDITIONS	% OF SAMPLE IN CROWDED CONDITIONS	% OF TOTAL TREATMENT SAMPLE
MT	100^{HI}	–	–	100
MTG	59	33	8	100
BT	84	16	–	100
BTG	40^{LO}	60^{HI}	–	100
BU	79	21	–	100
BUG	72	–	28^{HI}	100
TOTAL	434	130	36	600
% AND *MEAN*	(72%)	(22%)	(6%)	(100%)
RANGE	60	60	28	

CHI SQUARE = 253.90866 WITH 10 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

The floor conditions, i.e. the noise and crowding factors, which were present on the first floor of the museum, were noted at the start of each testing day and generally remained reasonably constant throughout the remainder of that day. Most of the visitors were tested under quiet conditions, when there were less than twenty people on the floor at any one time and when there were no people or objects, such as drills, making distracting noises. The two major exceptions were **BTG** and **MTG**, in which 60% and 41%, respectively, of the visitors were tested under noisy, crowded conditions. These two treatments also took the greatest number of visitors in the environment (the sample pool) in order to get the requisite tested sample of 100 people who looked at the display for at least 45 seconds. Three of the other treatments also support this premise that visitors in general spend less

time looking at a display under noisy, crowded conditions than they do in a quiet, uncrowded environment. These three treatments took a relatively greater population to achieve the requisite 45 second plus sample as the noise and crowding increased. The treatments were: **MT** with 0% noise and a sample pool of 434, **BT** with 16% noise and a sample pool of 447, and **BU** with 21% noise and a sample pool of 456. In addition, the two treatments which had no noise distractions (**MT** and **BUG**) also had the lowest sample pools (434 and 371 respectively).

Crowding was not generally a problem during the period that the treatments were tested, however when it did occur substantially, i.e. 28% in **BUG**, it did not seem to affect the short-term attention span of the casual visitor, since **BUG** had the lowest sample pool, 371. This might suggest that noise is a greater factor in affecting short-term attention spans (i.e. under 45 seconds) than is crowding. However it must be recognized that apart from some school parties, the crowding that was observed was minimal.

5.1.4. Size of Subject's Group Upon Entering the Floor:

All visitors who entered the gallery (excluding organized school groups) were observed and their group sizes recorded. The visiting group size was also recorded separately for those who took part in the tested portion of the experiment. Most of the tested subjects viewed the display treatment soon after entering the gallery and consequently their group sizes at entry were relatively easy to ascertain. Since there were generally so few groups of people in the gallery at any one time, and these groups often stayed together during their visit, it was also posssible to provide the entry groups for those tested visitors who looked at the display treatment long after they had first entered the floor.

TABLE 5.5- Number of tested subjects per treatment who entered the floor in each size of group:

DISPLAY TREATMENTS	ALONE	2 PEOPLE	3 OR MORE	ROW TOTAL
MT	36	49	15	100(16.7%)
MTG	25[LO]	57[HI]	18	100(16.7%)
BT	38	37[LO]	25[HI]	100(16.7%)
BTG	44[HI*]	39	17	100(16.7%)
BU	44[HI*]	41	15	100(16.7%)
BUG	35	56	9[LO]	100(16.7%)
TOTAL	222	279	99	600
% AND *MEAN*	(37%)	(46.5%)	(16.5%)	(100%)
RANGE	19	20	16	

CHI SQUARE = 23.16209 WITH 10 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0102

The great majority of tested subjects visited the floor either in two person groups (46.5%) or alone (37%). These proportions were similar to those for all other visitors to the floor. Overall only 16.5% of those tested arrived in groups of three or more. Proportions in the "three or more" group ranged from 25% in **BT** down to 9% in **BUG**. The chi square test shows a significant difference at the .01 level.

5.2. VISITOR RESPONSES TO DEMOGRAPHIC AND INTEREST QUESTIONS:

After completing the 'test' section of the questionnaire on the items and information in the display, the visitors were asked to answer ten short questions about themselves. These questions were aimed at finding out where they lived, whether or not they were frequent visitors to the museum, how they assessed their interest in and knowledge of Scottish archaeology, and their age, sex, occupation, and general educational level. Pilot studies indicated that these questions had to be kept short, general, and nonthreatening or they would not be answered at all. Therefore the basis for the subsequent demographic classifications was not a detailed questionnaire, it was a very simple set of questions which received general and even ambiguous replies, particularly about the subjects' occupations and educational levels.

There was a base core, of approximately 45 subjects in total, who completed little or none of the test or demographic sections of the questionnaire and who consequently accounted for many of the "missing value" responses. This base of constant "missing values" ranged from approximately five in **MTG** up to ten in **BT** and **BU**. The missing responses in general, however, seem to reflect the status value threat of the answers. The least number of missing observations, 58 and 59, occurred when the subjects were asked non-status questions about their past visits to the museum and their knowledge of and interest in Scottish history. More personal questions received fewer answers, i.e. residence with 64 missing values, age with 70, and occupation with 86 missing observations. Finally, the most status-threatening question, which asked about the subject's educational level, received the greatest number of null replies, 114.

5.2.1. (A. 1) Is This Your First Visit To This Museum?

5.2.2. (A. 2) Please estimate the number of times you've visited this museum in the last 12 months_____.

TABLE 5.6– Number of first time and repeat visitors in sample (within the repeat visitor group, those who visited recently, i.e. within the previous 12 months, have been additionally included in a separate column:)

DISPLAY TREATMENTS	FIRST TIME VISITORS	ALL REPEAT VISITORS	RECENT 1-6 TIMES	ROW TOTAL
MT	69[LO*]	22[HI*]	7	91 (16.8%)
MTG	83[HI]	12	8	95 (17.5%)
BT	79	8[LO]	6	87 (16.1%)
BTG	82	9	2	91 (16.8%)
BU	78	9	4	87 (16.1%)
BUG	69[LO*]	22[HI*]	12	91 (16.8%)
TOTAL	460	82	39	542
%	(85%)	(15%)	[7%]	(100%)
MEAN	77	14		90
RANGE	14	14		8

CHI SQUARE = 17.94377 WITH 5 DEGREES OF FREEDOM
for First Time and Total Repeat Visitors
SIGNIFICANCE = 0.0030

NUMBER OF MISSING OBSERVATIONS = 58

During this study, there were comparatively few repeat visitors to the museum who looked at the display treatments for a sufficient length of time to be tested. The largest number of repeat visitors, 22, occurred during the **MT** and **BUG** treatments. They represent a marked increase over the mean of 9 repeat visitors for the other display treatments. This surge could reflect Easter and summer holiday periods, when more local people might be expected to

visit.

Overall there was a significant association between the number of repeat visitors and the display treatments. Although the repeat visitor range was only 14, the significance level was .003. Therefore one might suspect the small variations in the numbers of repeat visitors between the Main Display Treatments of marginally influencing the evaluation measurement results. However, further statistical analysis showed that the differences in visitor attendance patterns between treatments were not significant below the .01 level, in their influence upon the questionnaire results (see Table 6.5).

The repeat visitors, as a whole, represented 15% of the sample tested. This percentage may be subdivided into recent repeat visitors, i.e, those who had visited the museum in the previous 12 months, and less recent repeat visitors. Each of these groups contained roughly equal numbers of visitors. When looked at in the context of the total sample, only 15% of the subjects tested had ever visited the museum before, including the surprisingly small number of frequent repeat visitors (7%) who had visited the museum at least once in the past year (39 out of the 542 who responded to this question). **BUG** had the largest number of recent repeat visitors, 12, and **BTG** had the smallest number, 2, closely followed by **BU** with 4. The remaining display treatments had similar numbers of repeat visitors in the past year, ranging from 6 to 8. The maximum number of times that one of the tested visitors said they had visited the museum in the last year was 6 times. Additionally, only 1% of the total had recently visited the museum more than two times, and the maximum number of recent repeat visits was only 6.

One might conclude, from table 5.6, that the museum had very few frequent repeat visitors and that the vast majority of visitors only came to the museum once. Alternatively, the repeat visitors may not have been interested in looking

at the introductory display case, and thereby avoided inclusion in the sample, preferring instead areas of special interest. The Control Group findings may be used to clarify this situation. The Control Group results for first time visitors (see Appendix G.II) suggest that during this experiment, when all visitors to the gallery were asked to complete a questionnaire, as in the ***Pilot Questionnaire (Unseen)***, ***Pre-Test*** and ***Ask to Study*** Groups, the vast majority (73%) were first time visitors. However the earliest Control Group, the ***Pilot Questionnaire (Unseen)***, which was tested in March, before the Main Display Treatments, had only 58% of its sample population who were first time visitors. Therefore, whereas it would appear that the number of repeat visitors was very low throughout the main portion of this study, these figures may be substantially increased over the winter months, and are consequently not necessarily representative of the visitor population over the entire year. Having said that, it must be remembered that the bulk of the museums' visitors come during the Spring and Summer months, and therefore may well consist of a large majority of first time visitors.

5.2.3. (B.) How interested are you in the early history and archaeology of Scotland?

TABLE 5.7– Sample's avowed interest in early Scottish History, per display treatment:

	INTEREST IN SCOTTISH HISTORY:			
DISPLAY TREATMENTS	VERY INTERESTED	MODERATELY INTERESTED	SLIGHT–NO INTEREST	ROW TOTAL
MT	17	46	29	92 (17.0%)
MTG	20^{HI}	40^{LO}	33^{HI}	93 (17.2%)
BT	15^{LO*}	51^{HI}	20^{LO}	86 (15.9%)
BTG	15^{LO*}	47	29	91 (16.8%)
BU	19	43	26	88 (16.3%)
BUG	18	43	30	91 (16.8%)
TOTAL	104	270	167	541
%	(19%)	(50%)	(31%)	(100%)
MEAN	17	45	28	
RANGE	5	11	13	

CHI SQUARE = 6.13575 WITH 10 DEGREES OF FREEDOM
SIGNIFICANCE = 0.8037

NUMBER OF MISSING OBSERVATIONS = 59

There was a fairly even spread of acknowledged interest levels, between the six display treatments, so it seems appropriate to discuss them as a whole. The significance level of .8 reinforces the similarity of the interested sample populations. Out of the 541 subjects who responded to this question, 19% reported themselves as being very interested, 50% as moderately interested, and 31% as slightly or not interested in Scotlands' early history and archaeology.

The similarity of interest levels, between the display treatments, particularly between those who were very interested in early Scottish history and archaeology, has beneficial repercussions for the analysis of the evaluative measurement results. One might have expected that any one treatment group which had an unusually large proportion of very interested visitors, would also have test scores, viewing times, etc. which were inflated by these motivated subjects. Fortunately, the homogeneity of the interest levels between the treatment samples, precludes the need to account for any differences in evaluation results which might have arisen from differences in interest levels.

5.2.4. (C.) How would you describe your general knowledge of Scottish (or British) archaeology?

TABLE 5.8- Sample's avowed knowledge of Scottish Archaeology, per display treatment:

DISPLAY TREATMENTS	POOR TO NO KNOWLEDGE	FAIR-VERY GOOD KNOWLEDGE	ROW TOTAL
MT	61	31^{HI}	92 (17%)
MTG	69^{HI}	25	94 (17.3%)
BT	56^{LO}	30	86 (15.9%)
BTG	68	24	92 (17%)
BU	67	20^{LO}	87 (16.1%)
BUG	66	25	91 (16.8%)
COLUMN TOTAL	387 (71%)	155 (29%)	542 (100%)
MEAN	65	26	
RANGE	13	11	

CHI SQUARE = 4.70061 WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE = 0.4535

NUMBER OF MISSING OBSERVATIONS = 58

This question was primarily asked in order to ascertain whether or not the tested subjects in any one treatment group were equipped with more previous knowledge of the subject, before they took the test, then were the subjects in of the other treatment groups. However, the chi square significance level of .45 shows that there was a fairly even spread of informed and ignorant subjects between the treatment groups. Taken as a whole, of the 542 who answered this question, 29% described themselves as having a fair to very good knowledge of British archaeology, and 71% professed themselves as having a poor or no knowledge of the subject. Consequently, the similarity of the results between the treatment groups suggests that the differences in the numbers of visitors who had a fair to very good knowledge of Scottish archaeology, were not sufficient to markedly influence the evaluation measurement results.

5.2.5. (D.) Where are you from?

TABLE 5.9– Sample's residence, per display treatment:

DISPLAY TREATMTS	EDINBRGH	SCOTLAND EXC.EDIN.	ENG,IRELD WALES	USA,CAN AUSTL,NZ	OTHERS	ROW TOTAL
MT	25^{HI}	20^{HI}	13	32^{LO}	1	91 (17%)
MTG	10	12	11	55	6	94 (17.5%)
BT	4^{LO}	10	12	56	5	87 (16.2%)
BTG	6	13	10	54	4	87 (16.2%)
BU	5	9	10	60^{HI}	3	87 (16.2%)
BUG	23	8^{LO}	13	40	6	90 (16.8%)
COLUMN TOTAL	73 (14%)	72 (13%)	69 (13%)	297 (55%)	25 (5%)	536 (100%)
MEAN	12	12	12	50	5	
RANGE	21	12	3	28	5	

6 OUT OF 30 (20.0%) OF THE VALID CELLS
HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 4.058

CHI SQUARE = 61.02673 WITH 20 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 64

Looking at the overall frequencies first, one interesting aspect was that over 60% of the subjects tested came from outside the United Kingdom and Ireland, mainly from the U.S.A. 536 of the subjects responded to this question. 14% were Edinburgh people and 13% lived in other parts of Scotland (making a Scottish total of 27%); 13% came from England, Wales, and Ireland; 55% were residents of other English speaking countries, i.e. the U.S.A., Canada, Australia, and New Zealand; and 5% came from non–English speaking countries. Therefore 95% of the subjects who answered this question came from English speaking countries.

Seasonal visitor changes appear to be reflected in residence changes in this sample population. There was a large upsurge in the number of Edinburgh visitors tested at the beginning and end of the study, from a mean of 6 for the four middle treatments, up to 25 for **MT** (Easter holidays) and 23 for **BUG** (beginning of the Summer holidays). The number from Scotland as a whole (Edinburgh plus the rest of Scotland) was also much higher in **MT**, when it was 28% of the combined Scottish total, but was fairly consistent in the other treatments, being 14% on average. The numbers in England, Wales, and Ireland, differed little from their average 16.5% (11.5 people). The U.S.A., etc group had markedly fewer people represented during **MT** (32) and **BTG** (40) than they did during the three middle treatments, where the average was 56 people. There were so few people in the group from non-English speaking countries, with a maximum of 6 people, that they are not worth discussing separately since all of these countries added together make-up less than 5% of the visitors tested.

The chi square significance level of .0 indicates a strong difference in residence locations between the treatment groups. Further analysis of the impact of these residence differences on the results of the questionnaire, however, showed that they were not significant below the .05 level (see Table 6.5).

5.2.6. (E.) Are you male or female?

TABLE 5.10- Sex of sample visitors, per display treatment:

DISPLAY TREATMENTS	MALE	FEMALE	ROW TOTAL
MT	48	52	100 (16.7%)
MTG	53	47	100 (16.7%)
BT	47^{LO}	53^{HI}	100 (16.7%)
BTG	50	50	100 (16.7%)
BU	51	49	100 (16.7%)
BUG	59^{HI}	41^{LO}	100 (16.7%)
TOTAL	308	292	600
% AND *MEAN*	(51%)	(49%)	(100%)
RANGE	12	12	

CHI SQUARE = 3.73599 WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE = 0.5880

Overall, 51% of the subjects tested were male and 49% female. These roughly equal numbers of males and females were generally consistent between the treatment groups, as is indicated by the 0.58 chi square significance level. The greatest variation occurred in **BUG**, where 59% of the subjects tested were male. In fact the proportions would have been 50/50 except for **BUG**'s high male population.

5.2.7. (F.) What is your occupation and what exactly do you do in it?

Subjects were classified by both socioeconomic group (O.P.C.S. 1970 HMSO Register of General Social Classes) and general occupational scale. However, because of the ambiguous nature of many of the responses, the socioeconomic groupings seemed prone to more errors than did the classification by general occupational scales, so the groupings have not been reproduced in table form.

Despite reservations as to their validity, it may be of interest to note the total percentages for the socioeconomic groupings. These were: professionals- 19% (96 subjects), managers- 30% (153), skilled- 10% (53), partly skilled- 2% (9), students- 17% (89), retired/unemployed/tourists- 10% (53), homemakers- 12% (61).

A simplified classification system was used for the general occupational scales. This included teachers, executives, and other people in responsible positions, in the professional category. The socioeconomic management classification was dispensed with entirely, since the visitors rarely provided enough detail about their jobs to form a true picture of their responsibilities. Other small categories, such as the 2% of unskilled workers, were added to the partly skilled, office workers, etc. to create the "Other" category. Homemakers were amalgamated with the retired, unemployed, and tourists to create an "Unwaged" category.

TABLE 5.11– Sample's broad occupational groups, per display treatment:

DISPLAY TREATMT	PROFES-SIONALS	STUDENTS	OTHERS-sales office,craft workers,etc	HOMEMAKERS RETIRED UNEMPLOYED	ROW TOTAL
MT	21^{LO}	20^{HI}	29^{HI}	20	90(17.5%)
MTG	39^{HI}	12^{LO*}	23	18	92(17.9%)
BT	31	12^{LO*}	16^{LO*}	19	78(15.2%)
BTG	24	17	22	18	81(15.8%)
BU	35	15	16^{LO*}	17	83(16.1%)
BUG	36	13	19	22	90(17.5%)
COLUMN TOTAL	186 (36%)	89 (17%)	125 (25%)	114 (22%)	514 (100%)
MEAN	31	15	21	19	
RANGE	18	8	13	5	

CHI SQUARE = 16.05870 WITH 15 DEGREES OF FREEDOM
SIGNIFICANCE = 0.3782

NUMBER OF MISSING OBSERVATIONS = 86

It was felt that those in the professional and student groupings were likely to be better trained to absorb knowledge quickly than were the subjects in the other categories. Consequently, major anomalies in the number of subjects in these two groups, between the display treatments, might have been expected to influence the test results. However, the number of subjects in each occupational group was fairly consistent between the treatments, as was shown by the 0.37 significance level. The anomalies which did occur were only minor ones, i.e.– 13% fewer professional people than average in **MT**, but 5% more students, and 8% more "others"; and 4% fewer "others" in **BT** and **BU**.

In this amalgamated general occupational scale, the largest category, of the 514 tested subjects who gave their occupations, was that of professionals, who

represented 36%. Students accounted for 17% of the tested population; and homemakers, the retired, and the unemployed accounted for 22%. "Other" occupations, which included office, restaurant, and hotel staff; craft workers; salespeople; farmers; etc, made up 25% of the occupations.

5.2.8. (G.) What is your age now?

Overall, 2% of the tested population fell into the "10 to 17" years of age bracket, 34% were between "18 to 29", 18% were between "30 to 39", 13% were between "40 to 49", 18% were between "50 to 59", and 15% were between "60 to 83". The "18 to 29" age band obviously represented the major group of tested subjects. It must be remembered, however, that visitors under the age of 14 were discouraged from completing a questionnaire and only did so with the assistance of an adult. (For an indication of the number of children present in the gallery population, see Appendix L.) 530 of the potential 600 subjects answered this question.

TABLE 5.12- Sample's age groups, per display treatment:

DISPLAY TREATMT	10-17	18-28	30-49	50-83	ROW TOTAL
MT	3	29	26	30	88(16.6%)
MTG	4	30	35^{HI}	25	94(17.7%)
BT	2	29	16^{LO}	39^{HI}	86(16.2%)
BTG	2	32	27	26	87(16.4%)
BU	3	27	33	23^{LO}	86(16.2%)
BUG	0	31	26	32	89(16.8%)
COLUMN TOTAL	14 (3%)	178 (33%)	163 (31%)	175 (33%)	530 (100%)
MEAN	2	30	27	29	
RANGE	4	5	19	16	

6 OUT OF 24 (25.0%) OF THE VALID CELLS
HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 2.272

CHI SQUARE = 18.12419 WITH 15 DEGREES OF FREEDOM
SIGNIFICANCE = 0.2562

NUMBER OF MISSING OBSERVATIONS = 70

By amalgamating the age groups even further, as in Table 5.12, it would be possible to assert that roughly one-third of the subjects tested fell into each of the following age bands: "18 to 29", "30 to 49", and "50 to 83". The number of subjects in each age group varied little between display treatments, particularly in the "18 to 29" age band, and only varied more than 10 subjects from the group mean in **BT**. In the "30 to 49" year old group, **BT** has a sample population of only 16, which is 12 subjects below the mean of 27, and in the "50 to 83" year old group, **BT** rises to 39, which is 11 more than the group mean of 29. However the chi square significance level of 0.25 suggests that the differences which did occur were not significantly important.

5.2.9. (H.) At what age did you stop being a full-time student?

TABLE 5.13- Samples' ages when their full-time education stopped, per display treatment:

DISPLAY TREATMT	AGE WHEN EDUCATION STOPPED: 14-16 + 17-18		19-22 + 23-89 DEGREE? POSTGD?		STUDENT NOW	ROW TOTAL
MT	19 + 5=24^{HI}		17 + 20=37^{LO}		20^{HI}	81(16.7%)
MTG	9 9=18		28 31=59^{HI}		12^{LO}*	89(18.3%)
BT	5 + 4= 9^{LO}		25 + 27=52		12^{LO}*	73(15.0%)
BTG	9 8=17		31 19=50		17	84(17.3%)
BU	6 + 9=15		20 + 30=50		15	80(16.5%)
BUG	11 8=19		26 20=46		14	79(16.3%)
COLUMN TOTAL	59 (12%)	43 (9%)	147 (30%)	147 (30%)	90 (19%)	486 (100%)
MEAN		=17		=49	15	
RANGE		=15		=22	8	

CHI SQUARE = 29.03188 WITH 20 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0871

NUMBER OF MISSING OBSERVATIONS = 114 (19%)

12% of the tested population said that they stopped being a student between the ages of "14 to 16", 8% between "17 to 18", 30% between "19 to 22", 30% between "23 to 89", and 19% were students now. These groupings can be simplified even further by stating that 19% of the visitors tested were students, 20% had completed their full-time studies by the time they were 18, and 60% had continued their education after the age of 18. However since 19% of the visitors tested did not answer this question, it may be that many of them would have fallen into the "before 18" group and were ashamed to admit this on their questionnaires.

Comparing the display treatments between groups, most were reasonably

similar, with the greatest range being 22 in the "over 18" age group. **MT** portrayed the greatest deviation from the means of all three groups. It had more subjects in the "18 and under" category than did any other display treatment (24, which is 7 more than the group mean of 17), more students (20, which is 5 more than the group mean of 15), and inevitably, fewer in the "over 18" category (37, which is 12 below the category mean of 49). **MTG** had the largest number of subjects in the "over 18" group (59, which is 10 more than the group mean of 49, and 7 more than its nearest rival). **BT** had the least number in the "18 and under" category. It had 9 subjects, which is 8 below the group mean of 17, and 6 below its nearest competitor. However, the age groups demonstrated a chi square significance of 0.087, which, although close to the demarcated significance level, in this study, of 0.05, manages to be far enough away so that these groups are not considered significantly different from one another.

5.2.10. (I.) Did you know you were going to be asked to fill-out this questionnaire?

TABLE 5.14- Samples' avowed prior knowledge of questionnaire, per display treatment:

DISPLAY TREATMENTS	DIDN'T KNOW ABOUT TEST	KNEW ABOUT TEST	ROW TOTAL
MT	90	1	91 (17.1%)
MTG	92	0	92 (17.3%)
BT	84	0	84 (15.8%)
BTG	90	1	91 (17.1%)
BU	88	0	88 (16.5%)
BUG	87	0	87 (16.3%)
COLUMN TOTAL	531 (99.6%)	2 (0.4%)	533 (100%)

6 OUT OF 12 (50.0%) OF THE VALID CELLS
HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.315

CHI SQUARE = 3.87167 WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE = 0.5680

NUMBER OF MISSING OBSERVATIONS = 67

Only two subjects out of the 533 who responded to this question said that they had guessed that they might be asked to fill in a questionnaire. Consequently almost none of the visitors admitted to any prior knowledge that they might be asked to complete a test, and could not, therefore, have had their habits influenced by this knowledge. (Assuming, of course, that they answered this question truthfully.)

5.3. SUMMARY

5.3.1. General Chapter Summary:

Great variety was allowed to occur between the display treatments in the number of days taken to collect a sample, the days of the week which were tested, and the relative sizes of the weekend versus the weekday sample populations. Correlations were found between the number of days needed to obtain the sample quota, the number of hours used in observing and testing, and the proportion of subjects drawn from small sample pools (the latter being a reverse correlation). It was also found that visitors were less likely to look at a display for 45 seconds or more in a noisy environment than in a quiet one.

Over half of the subjects were tested on days which began with sunshine and only 11% with rain. Most were also tested under quiet, uncrowded conditions, and visited the gallery either with one other person or alone.

There appeared to be a relationship between the number of null responses and the status-threatening value of the answers. The questions on educational level attained, and on occupation, received far more than the usual number of null replys.

The greatest number of local and repeat visitors came during Easter and Summer holiday periods. Overall, there were few repeat visitors tested. It may be that the museum actually attracts few frequent visitors, or that the frequent visitors are not likely to be attracted to an introductory case, preferring instead areas of special interest. (However since quite a few of the repeat visitors had not visited the museum in the past year, it may be that few actually do return frequently.) An obvious reason for the lack of repeat visitors is that most came from abroad, primarily the U.S.A. Only 40% of those tested came from the

United Kingdom, including the small proportion (14%) from Edinburgh.

The levels of visitor interest in Scottish history and archaeology, and their professed knowledge in this area, were very similar across all of the treatments. Less than a third of the subjects felt they had a fair to very good knowledge of these topics, and only a fifth stated that they were very interested in these areas. Perhaps the scarcity of very interested and knowledgeable visitors was due to the lack of repeat visitors.

For most of the subjects tested, it was their first visit to the museum, and they were moderately interested in early Scottish history and archaeology, but felt they had little knowledge in these areas. In fact the typical subject would have been an American professional person, in his or her twenties, who had continued in full-time education beyond the age of 18, and who was unaware that experimental testing was being conducted.

The interest and knowledge levels, and the sex of the visitor samples, were very similar. The maximum range in any one category was 13. The age groups were also very similar, with a maximum range of only 7, provided that the two groups of over 30 year olds are merged. Although the education levels and occupational groups were not as similar as the other groups mentioned, they were still well above the 0.5 chi square significance level.

Differences were noted in the number of repeat visitors, the test conditions, and the residence of visitors. Although the differences between the number of repeat visitors was relatively small, with a range of only 14, the means for two of the treatments were distinctly separated from those of the other four treatments. However, in Chapter 6, following Table 6.5, a discussion is presented which demonstrates that the differences which did exist were not likely to have influenced the evaluation measurement results.

5.3.2. Attendance Implications for All Museum Visitors:

It would seem reasonable to suggest that the visitors observed during this study were representative of the general visitors to the museum at the time. During this study, the large gallery on the ground floor was temporarily closed for renovation, so all visitors were directed by sign up the stairs to the two other floors. The First Floor Gallery, in which this study was undertaken, would consequently be the first gallery visitors would be able to enter after coming into the museum. Several studies (O'Hare, 1974) have shown that visitors see proportionally fewer general galleries, the further they are from the main entrance. Therefore it would seem reasonable to conclude that the first gallery would receive the greatest representative sample of visitors. The Roman Gallery, on the second floor, was moreover comparitively small, and it would therefore also seem reasonable to assume that most of those visitors who first visited the Roman Gallery, would also look around the other gallery, on the first floor.

A complication, which affects conclusions about visitor attendance figures, is the physical situation of the museum. It shares a building with the Scottish National Portrait Gallery, each institution occupying different sides of the building, with shared corridor access on the floors. Therefore all official attendance figures are necessarily joint ones, for both museums, and it is consequently impossible to accurately assess the separate attendance figures for the Museum and the Gallery.

It therefore remains important to establish if the tested visitors' sample was representative of all of the visitors to the First Floor Gallery, and consequentially, to the Museum. In the few cases, such as Sex, where comparative figures are available for both the tested visitor sample and the total visitor population, the populations do appear to be similar (see Appendix

L). Likewise, comparisons with Pilot and Control Groups, such as the ***Pilot Questionnaire*** Study, the ***Ask to Study,*** and the ***Pre-Post*** Groups, all of which solicited the co-operation of all adult visitors who entered the Floor, indicate that the tested visitor sample was representative of the available visitor population (see Appendices K,L,M).

5.3.3. Implications for Comparison with Other Studies:

The demographic characteristics of the general visitors to the National Museum of Antiquites of Scotland (N.M.A.S.) are not fully known. However the sample in this study appears to be representative of the visitors who came to the museum between April and July. Consequently, tenuous comparisons may be made with other studies, when similar data exists. In the following tables, the initials "ROM" refer to the Royal Ontario Museum (Abbey and Cameron, 1959-61), and "AMNH", to the American Museum of Natural History (1977), in New York City. The designation "British Ms" is the British Museum, in London; "Science Ms" is the Science Museum, in London; "Maritime Ms" is the Maritime Museum, in Greenwich (Wingfield Digby, 1974); "B.M.Nat.His." is the British Museum of Natural History, in London (Alt, 1978); and the final two are the Manchester (Mason, 1974) and Norwich Castle (McWilliams and Hopwood, 1973) Museums.

5.3.3.1. Group size–

MUSEUM	ALONE	2 PEOPLE	3 OR MORE
N.M.A.S.*	**37%**	**46.5%**	**16.5%**
ROM	31%	36%	33%
AMNH	13%	n/a	n/a
BRITISH MS.	28%	n/a	n/a
SCIENCE MS.	13%	n/a	n/a
MARITIME MS.	7%	n/a	n/a
B.M.NAT.HIS.*	31%	37%	32%exc sch gp
MANCHESTER	28%	n/a	n/a
NORWICH CASTLE	14%	34%	52%(incl sch gp)

*– these starred museums excluded school groups from their calculations.

The N.M.A.S. Spring and early Summer sample had a higher proportion of visitors who came alone than did any of the other museums listed. However many of these museums included school and other organized groups in their findings. Consequently their "3 or More" category results are inflated and their "Alone" results are correspondingly deflated, when compared with those who did not include organized parties. The Royal Ontario Museum and the British Museum of Natural History had the closest proportion of solo visitors, 31%, to the 37% in the N.M.A.S. group.

5.3.3.2. First Timers-

MUSEUM:	FIRST TIME VISITORS	ALL REPEAT VISITORS	(RECENT) (VISITORS)
N.M.A.S.	**85%**	**15%**	7
ROM	24%	76%	n/a
AMNH	20%	80%	40%
BRITISH MS	47%	53%	34%
SCIENCE MS	46%	54%	29%
MARITIME MS	72%	28%	15%
B.M.NAT.HIS.	49%	51%	19%
MANCHESTER	24%	76%	53%
NORWICH CASTLE	39%	61%	36%-last 6 mo.only

The proportion of first time visitors to the N.M.A.S. was unusually high, when compared with the other museums listed above. This may be because local winter visitors were not accounted for in this sample. However, the results from the National Maritime Museum also show an unusually high proportion of first time visitors.

5.3.3.3. Residence-

MUSEUM	LOCAL CITY	COUNTRY	ABROAD
N.M.A.S.	**13%**	**26%(U.K.)**	**60%**
ROM	n/a	62% incl.city	38%
AMNH	42%	54%	4%
BRITISH MS	35%	32%	33%
SCIENCE	38%	50%	12%
MARITIME	41%	42%	17%
B.M.NAT.HIS.	28%	32%	40%
MANCHESTER	84%	16% all others	n/a
NORWICH CASTLE	23%	70%	7%

In this Springtime and early Summer study, the N.M.A.S. attracted an exceptionally low proportion of visitors from the surrrounding city, and a very high proportion from abroad.

5.3.3.4. Sex-

MUSEUM:	MALE	FEMALE
N.M.A.S.	**51%**	**49%**
ROM	51%	49%
AMNH	54%	46%
BRITISH MS	53%	47%
SCIENCE MS	62%	38%
MARITIME MS	59%	41%
B.M.NAT.HIS.	59%	41%

In this study the N.M.A.S. appeared to attract a slightly higher proportion of female visitors then did most of the other museums mentioned.

5.3.4. Main Differences Between the Six Display Treatments:

MT took close to the average number of days and hours for testing and had a below average sample pool. However it had the highest weekend sample population, which was almost twice the size of the mean. While it had an average number of visitors from "small sample pool" groups, it had a very large number from "large sample pool" groups. The weather during this display treatment was largely sunny and only a few of the subjects experienced a rainy start to their days. This was also the only treatment group to experience entirely quiet, uncrowded floor conditions. Almost half of the subjects in this group arrived with one other person.

This display treatment group contained one of the highest proportions of repeat subjects but only an average number of frequent repeat subjects. The subjects expressed average amounts of interest in Scottish History and had the highest proportion of self-assessed, fair to very good, knowledge of the subject. This treatment group also contained the highest proportion of Edinburgh and Scottish residents, the highest U.K. total, and the lowest from outside the U.K.. It also contained the largest number of students and

office-sales-craft workers etc., the smallest number of professionals, and a high number of homemakers-retired-unemployed. This sample's age and sex were close to the mean for these groups but they had the highest proportion of subjects whose full-time education had been completed by the time they were 18.

MTG took more than the average number of hours and days for testing and had the largest sample pool. However its weekend sample population size was close to the average. It had the largest number of subjects from "small sample pool" groups and an average number from "large sample pool" groups. **MTG** had the lowest proportion of sun and the highest of rain. The subjects also faced a fairly high proportion of noisy initial conditions and a small proportion of crowded conditions. This treatment group had the highest proportion of subjects who entered the gallery with one other person and the lowest who arrived alone.

MTG had the highest proportion of subjects who were first time visitors to the museum but a slightly above average number of frequent repeat visitors. It also had the highest proportions of subjects who were very interested, and who had slight to no interest, in early Scottish History, and who felt they had little knowledge of the subject. The subjects in this treatment group had reasonably average residence and sex profiles and the highest proportion of professionals, 30 to 49 year olds, and subjects who continued their studies beyond the age of 18. It also contained one of the lowest proportions of students and "null" responses.

BT had a fairly average sample pool size and number of hours and days to complete its sample quota. It also contained, by far, the smallest number of subjects from a weekend population. It had a very large number of visitors from "small sample pool" groups, but none from a "large sample pool". **BT** also

had the highest proportion of clouds, but no rain, and a high proportion of quiet conditions and no crowded starts. It had the highest number of subjects who entered the floor in groups of three or more people but the lowest who arrived with one other person.

This treatment group had the lowest proportion of repeat visitors to the museum in its sample but an average number of them were frequent repeat visitors. It also had low proportions of subjects who were very interested in Scottish History and those who had little interest or knowledge of the subject. It had the lowest number of Edinburgh residents, average numbers from the rest of the U.K., and a fairly high proportion from the U.S.A. etc. It also had the highest proportion of females and of people between the ages of 50 to 83. It had one of the lowest numbers of students and of office, sales, etc. workers and the lowest for the 30 to 49 age range and for those who had stopped their education by the age of 18.

BTG had a high number of visitors present in its sample pool and took the largest number of days and hours to complete the testing. It also had a very low number of subjects in its weekend population. It had a high number of subjects from "small sample pool" groups, but none from "large sample pools". It also had the lowest proportion of cloudy and quiet starts to a day, no rainy or crowded starts, and the highest proportion of noisy beginnings. It had one of the highest proportion of subjects who entered the gallery alone.

This treatment group had a very low proportion of repeat visitors in its sample and the lowest number of frequent repeat visitors. It also had a very low proportion of subjects who were very interested in Scottish History. It had a reasonably low proportion of subjects who were from Edinburgh and who were in the professions but average male-female, age, and education level groups.

BU took a below average number of hours and days for testing and had an average number of visitors in both its sample pool and its weekend sample. It had a very low number of subjects from "small sample pools" and the highest number from "large sample pools". The subjects in **BU** faced mixed beginnings to their days. Almost half experienced cloudy weather, a third sun, and a fifth rain. However most met with quiet conditions and none faced crowded ones. This treatment group also contained one of the highest proportions of subjects who came to the gallery alone.

BU had a very low proportion of both repeat, and frequent repeat, visitors in its sample. It also had the lowest proportion of subjects who felt they had much knowledge of Scottish History. It had the highest proportion of subjects from the U.S.A. etc. and one of the highest in the 30 to 49 age range. It had one of the lowest proportions of U.K. residents, of sales-office-etc. workers, and of 50 to 83 year olds. Its sex and education level proportions were, however, close to the average.

BUG took the least number of hours and days for testing and had the smallest available sample pool. It also had a relatively large number of subjects in its weekend sample. It had the lowest number of visitors from "small sample pools" and a fairly average proportion from "large sample pools". The subjects in this treatment group experienced a high proportion of daily sunny beginnings, no rainy ones, and only a moderate amount of cloudy starts. They also faced the highest proportion of initially crowded conditions, but the majority had quiet starts. It had a very high proportion of subjects who entered the gallery with one other person, and the lowest number who entered with more than one other person.

This treatment had one of the highest proportions of repeat, and frequent repeat, visitors in its sample. Its subjects expressed average interest in early

Scottish History and knowledge of that subject. A high proportion of the subjects came from Edinburgh, were male, and were professionals. Low numbers came from the rest of Scotland and from the U.S.A. etc. There were average numbers of subjects in the different age and education level groups.

CHAPTER 6

RESULTS OF THE DEPENDENT VARIABLES:

INITIAL ATTRACTION, INTEREST RETENTION, KNOWLEDGE GAIN, WRITTEN COMMENTS

Four factors have been selected to provide a quantifiable measure of the communication effectiveness of each of the display treatments. A comparison of these measures between the display groups may then point to the strong features of each, thereby enabling an informed choice to be made in selecting one of these display formats for use in other exhibition environments.

The four measures of communication effectiveness are: *Initial Attraction*, *Interest Retention*, *Knowledge Gain*, and *Written Comments*. The *Initial Attraction* measurement reflects the power each display treatment had to capture the visitors' attention when they first entered the gallery. *Interest Retention* demonstrates the length of time each treatment was able to keep the visitors' attention. *Knowledge Gain* reflects the amount of information visitors' could recall after viewing a display. The *Written Comments* suggest whether visitors generally had a favourable or unfavourable opinion of the display treatment which they saw. Both the *Knowledge Gain* and the *Written Comment* factors apply only to the sample of visitors who were asked to complete the written questionnaire. The *Initial Attraction* and the *Interest Retention* factors, on the other hand, reflect the observed attitudes of all of the visitors who were exposed to the display treatments during the study.

Whereas the emphasis in the previous chapter was on accounting for any differences in test conditions and demographic characteristics which might have biased the results of a particular display treatment, the concern in this chapter is to take an overview of the treatment evaluations. This chapter sets out to compare visitor reactions to different combinations of treatments in

order to determine the "best" display treatment conditions. These primary evaluative measures are expressed in terms which reflect the museum visitors' responses towards the exhibit case. If the visitors' responses to a particular display format (as expressed in one of the four primary evaluative measures) are significantly greater than the quantifiable responses to one of the other treatments, then that display treatment is deemed to be better at eliciting the response in question.

The first part of this chapter looks at each of the four evaluative communication factors in turn, i.e. *Initial Attraction, Interest Retention, Knowledge Gain*, and *Written Comments*. Within each of them, the six Main Display Treatment results are assessed individually, as well as being amalgamated into larger units for additional comparisons. These amalgamated units are: 1) the Principal Display Treatments (**MTs, BTs, BUs**); and 2) the treatments *With Goals* (**MTG, BTG, BUG**) and *Without Goals* (**MT, BT, BU**). The major differences, and similarities, between each of these groups will be discussed.

The second section of the chapter demonstrates, in a naturalistic enquiry fashion, the range of written comments expressed by the visitors. In the last section, the chapter again examines each one of the four evaluation factors, and compares the means for the Main Display Treatments with the relevant Control Group results. The chapter concludes with a summary and an overview of the communication effectiveness of the display treatments.

1. EXPECTATIONS FOR THE THREE PRINCIPAL DISPLAY TREATMENTS AND FOR GOALS

One would expect the Multi-Track Labels and Objects Treatments (**MTs**) to emphasize themes and the evolution of ideas through the Ages, with slightly

less emphasis on chronological development within the Ages. The label format, with its clear bold headings, should make it easier for people to find the information they seek and to get an overview of the contents of the display case. One might also expect that since the primary label content is more accessible, being broken-up into short paragraphs, the visitors might read and remember more details. Grouping the objects into categories should also make it easier for the visitors to find the items which interest them and to observe their evolution through time. These object groupings should also enable the interested visitor to gain a clearer understanding of the differences in general life-styles between the Ages. One would consequently expect many positive comments about the clear, logical display. However one would also expect many negative comments against the crowding of objects. (Slight crowding was necessary because of the need to leave sufficient blank space between the object groupings inorder to separate them.) One might expect the best overall performance from this treatment.

From the Block Labels and Multi-Track Objects Treatments (**BTs**) one might expect an emphasis on visual observation of objects and their evolution and developments across the Ages. Headings and object groupings still emphasize themes but one might expect that most visitors would absorb general rather than specific information, since the headings are separated from their relevant descriptive paragraphs. One might expect the worst performance from this display since its objects and labels are arranged in different formats.

From the Block Labels and Unstructured Object Arrangement (**BUs**) one might expect much less emphasis on themes and the evolution of traditions across the Ages since it is harder to follow a visual connection (except with the Burials group which was left at the bottom of the display for all of the treatments) between the unstructured objects and labels in each Age. Instead, one would expect an emphasis on Ages and life-styles. Fewer crowded

complaints might also be anticipated since the objects were more spaced out. One might anticipate a reasonable performance from this typical type of museum display.

Goals (..G)– when goals were added to the above display treatments, one would expect that they would emphasize the evolution of the four themes across the Ages and stimulate greater interest and learning. One possible negative effect however is that they made the displays appear rather fussy and consequently, less attractive, by introducing extra written information, colour, and lines.

6.1. ATTRACTION:

The *Initial Attraction* value for each display treatment was determined by dividing the total number of visitors to the hall by the number who initially went straight to the display case, during the time that the particular treatment was under study. The result is thus a percentage figure which shows the ratio between the number of visitors who entered the environment and the number who first looked at the display case, before viewing anything else in the hall. The measurements were obtained by direct observation of the visitors' traffic patterns as they entered the hall.

TABLE 6.1- Initial Attraction to the Main Display Treatments:

DISPLAY TREATMENTS	% OF INITIAL ATTRACTION	TOTAL NO.OF VISITORS	NO.OF INITIAL VIEWERS
Principal:			
MT+MTG	32.7%	939	307
BT+BTG	30.3%	947	287
BU+BUG	33.9%	827	281
Goals:			
WITH GOALS	32.1	1376	442
WITHOUT GOALS	32.4	1337	433
Main:			
MT	35.7%	434	155
MTG	30.1%	505	152
BT	30.6%	447	137
BTG	30.0%^{LO}	500	150
BU	30.9%	456	141
BUG	37.7%^{HI}	371	140
MEAN	32.5%	452	146
RANGE	7.7%	134	18

CHI SQUARE= 10.549 WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE= GREATER THAN .05 AND LESS THAN .1

The Main Display Treatments, i.e. **MT–BUG**, generally varied little from their mean *Initial Attraction* value of 32.52%, with a range of less than 8%. As shown by the chi square test, the differences between treatments were not significant below the .05 level. The differences were further minimized when these treatments were amalgamated into the three Principal Groups, by adding the basic treatment to its relevant goal treatment, ex: **MTs= MT+MTG**. The resultant means ranged from 34.33% for the **BUs** down to 30.32% for the **BTs**. There was also little difference between the means of the treatments *With Goals* and those *Without Goals*.

6.2. INTEREST RETENTION:

The *Interest Retention* measurement records the average length of time that all visitors looked at the display case, i.e. the total viewing (and glancing) times divided by the total number of viewers. The overall results of this measurement were very clear cut. The results within the test treatments were markedly higher when goal orientating devices were present. (For example, the *Interest Retention* time was higher in **BUG**, which contained goals, than in **BU**, which was the same display treatment without goals.) However because of the way viewing times were coded, so that they were aggregated for each treatment, as opposed to being recorded for each individual, it was not possible to carry out significance tests on this data.

TABLE 6.2- Interest Retention time, in seconds, for the Main Display Treatments:

DISPLAY TREATMENTS	MEAN TOTAL INTEREST TIME	TOTAL TIME	TOTAL NUMBER OF VIEWERS
Principal:			
MT+MTG	120 seconds	59174"	492
BT+BTG	93 seconds	53137"	569
BU+BUG	129 seconds	56749"	440
Goals:			
WITH GOALS	122 seconds	90639"	741
WITHOUT GOALS	103 seconds	78421"	760
Main:			
MT	114 seconds	28722"	252
MTG	127 seconds	30452"	240
BT	82 seconds^{LO}	20807"^{LO}	255
BTG	103 seconds	32330"^{HI}	314^{HI}
BU	114 seconds	28892"	253
BUG	149 seconds^{HI}	27857"	187^{LO}
MEAN	113 seconds	169060"	1501
RANGE	67 seconds	11523"	127

As if to further reinforce the importance of good design techniques in retaining the viewer's interest, within the Principal test Treatments, the rather internally inconsistent **BTs** displays scored lower in their *Interest Retention* times than did the more cohesive **MTs** and **BUs** display treatments. The **BTs** display, with its primary information contained in a single information "block" but its artifacts arranged in separate "tracks", had a mean total interest time of 93 seconds. The **MTs**, or "multi-track" treatments, which divided both objects and labels into separate, compatible, "tracked" groupings which were consistent across the twelve foot wide display case, had a mean interest time of 120 seconds (two minutes). The **BUs** displays with their "block" information labels and "unstructured" object placement, had a mean *Interest Retention* time of 129 seconds (two minutes and twelve seconds). This is 7% higher than the **MTs** total. However, since both **MT** and **BU** alone had the same *Interest Retention* score, it may be that the increase is primarily due to the influence of the goal statements.

As a group, the display treatments *Without Goals* kept the viewer's interest for an average of 103 seconds (one minute and 43 seconds) while the same displays, *With Goal* orientating material, increased the average interest by 19 seconds (16%) to 122 seconds (two minutes and two seconds). The distinction between the *With Goal* versus the *Without Goal* displays was even more pronounced when one looks only at the internally consistent **MTs** and **BUs** groups. If these groups alone are amalgamated, the use of *Goals* increased the average viewer's interest span by 24 seconds, from 114 seconds (one minute and fifty-four seconds), to 137 seconds (two minutes and seventeen seconds). Therefore the use of *Goals* would appear to markedly increase the amount of time that visitors spent looking at the display treatments.

6.3. KNOWLEDGE GAIN:

The *Knowledge Gain* measurement was derived from the results of a test. This test formed the larger part of a two page questionnaire. In calculating the test scores, each of the eighteen test items were given equal weight, with the obvious result that "18" was the highest score possible.

TABLE 6.3- Knowledge Gain scores for the Main Display Treatments:

DISPLAY TREATMENTS	KNOWLEDGE GAIN MEAN TOTAL TEST SCORES(max= 18)	HIGHEST SCORE (max=18)	NO.OF NULL SCORES
Principal:			
MT+MTG	5.14 (28.6%)	14.5 (mean)	8 (mean and %)
BT+BTG	4.33 (24.1%)	14 (mean)	10.5 (mean and %)
BU+BUG	4.85 (26.9%)	14 (mean)	14 (mean and %)
Goal:			
WITH GOALS	4.95 (27.5%)	16 (mean)	8 (mean and %)
WITHOUT GOALS	4.60 (25.6%)	12 (mean)	13 (mean and %)
Main:			
MT	5.07 (28.2%)	13	9
MTG	5.20 (28.9%)^{HI}	16	7
BT	4.01 (22.3%)^{LO}	10	14
BTG	4.65 (25.8%)	18	7
BU	4.71 (26.2%)	14	17
BUG	5 (27.8%)	14	11
MEAN	4.77	14	11
RANGE	1.19	8	10

Number of Missing Observations (Null Scores) = 65

ANALYSIS OF VARIANCE:
Knowledge Gain Scores for the Main Display Treatments

DISPLAY TREATMENTS	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
Between Treat.	92.336	5	18.467	1.818
Within Treat.	6034.437	594	10.159	
Total	6126.773	599		

SIGNIFICANCE= 0.107

There was some difference between the mean Test Scores for the Principal Display Treatments, i.e. **MTs, BTs, BUs**. The **MTs** were the highest, with 28.6% correct answers. The **BUs** were next with 26.9%, and the **BTs** were the lowest with 24.1% of their answers correct. There was however, little to distinguish between the mean High Scores (highest scores attained) of the Principal Treatments. The **MTs** had a mean High Score of 14.5, out of a possible 18, closely followed by the **BTs** and **BUs** mean High Score of 14. The mean Null Scores, on the other hand, did show some differences. The Null Score, in this case, represents the number of "0" (zero) test scores. Visitors obtained a Null Score either by correctly answering none of the test questions or by returning a blank questionnaire. The **MTs** displayed the lowest number of Null Scores, with a mean of 8 (8%). The **BTs** had a mean of 10.5% and the **BUs** had the highest number of zero and blank responses, 14%. Looking at the individual Main Display Treatments, all of the treatment groups demonstrated increases in their mean Test Scores when *Goals* were present. The High Scores also increased or remained the same under the presence of *Goals*, while the number of Null Scores decreased. These results suggest therefore that the **MTs** in particular, and *Goals* in general, had a slightly beneficial effect on Test Scores, High Scores, and Null Scores.

A conservative test of the analysis of variance, conducted across the six Main Treatment groups, did not yield significant results for *Knowledge Gain*, as the previous table demonstrates. However when the analysis of variance was

conducted separately on the three Principal Treatments, and on the treatments *With Goals* and *Without Goals*, some significant results did emerge. Whereas there was not a significant difference between the treatments *With Goals* and those *Without Goals*, the differences in *Knowledge Gain* between the Principal Display Treatments were significant at the .05 level (DF=2, F=3.291). However in view of the failure to uncover significant differences in *Knowledge Gain* across the six Main Display Treatments, the significant differences which occurred between the Principal Display Treatments must necessarily be viewed as equivocal, and inferences must therefore be cautious.

6.4. WRITTEN COMMENTS:

All of the visitors who were given a questionnaire received a written invitation, at the end of the questionnaire, to add any additional comments of their own in the space provided, or to continue onto the back of the page. These free format comments were broadly classified into comment groupings and also classified as being either positive or negative responses. The general comment groupings which were used are listed in section 6.4.1.1. Subsequent sections present a representative selection of the range of comments which occurred in each group. For examples of positive and negative classification, see section 6.4.1. The first comment in that section was classified as being positive and the second comment as being negative. The final comment in that section was classified as being both a negative design comment and a positive label comment.

In order to carry out this dual classification without bias to any one treatment group, each comment was placed on a separate index card and its relevant treatment group and questionnaire number were placed out of sight on the opposite side of the index card. Therefore the classification could not be influenced by any preconceptions as to the optimum treatment. Since the

same person was responsible for assigning all of the the visitors' comments into positive, negative, and group categories, and did so "blind", any biases which did exist were uniform across the treatment groups.

The *Written Comments* were coded, for one set of calculations, so that any number of positive responses by a subject would count as only one positive respondent result, and the same for the negative responses. Obviously, however, many of the subjects wrote both positive and negative comments, and under this system such a visitor would be coded as both a positive and a negative respondent. In the following table, the total number of negative respondents were subtracted from the total number of positive respondents who commented on that display, to yield the true, or "resultant", number of positive or negative respondents.

In another calculation, the total number of negative comments were subtracted from the total number of positive comments, to yield the "resultant" number of positive or negative comments. The number of subjects who did not either return their completed questionnaires or write down any comments have also been calculated for each display group. They are noted under the heading "Number of Missing Observations".

TABLE 6.4- Number of positive minus negative respondents and comments for the Main Display Treatments:

DISPLAY TREATMENTS	RESULTANT RESPONDENTS	RESULTANT COMMENTS
Pricipal:		
MT+MTG	46 positive (23%)	357 positive
BT+BTG	3 positive (1.5%)	237 positive
BU+BUG	13 positive (6.5%)	261 positive
Goals:		
WITH GOALS	39 positive (13%)	374 positive
WITHOUT GOALS	23 positive (7.7%)	485 positive
Main:		
MT	21 positive	194 positive^{HI}
MTG	25 positive^{HI}	163 positive
BT	3 negative^{LO}	132 positive
BTG	6 positive	105 positive^{LO}
BU	5 positive	159 positive
BUG	8 positive	106 positive
MEAN	10 positive	143 positive
RANGE	28	89

Number of Missing Observations = 35

ANALYSIS OF VARIANCE:
Resultant Respondents for the Main Display Treatments.

DISPLAY TREATMENTS	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
Between Treat.	69.895	5	13.979	2.141
Within Treat.	3649.056	559	6.528	
Total	3718.952	564	6.594	

SIGNIFICANCE= 0.059

23% of the 200 subjects who were tested for the **MTs** Treatments were resultant positive respondents. On the other hand, only 1.5% of the **BTs**

subjects, and 6.5% of the **BUs** subjects, were resultant positive respondents. This remarkable lack of resultant positive respondents could have been due to seasonally different populations in the later treatments. However the **BUG** visitor characteristics are similar to those of **MT**, both having a high proportion of Edinburgh residents (see Table 5.9). Therefore it would appear that an overwhelming majority of the subjects were positive in their assessments of the **MTs** Treatments when compared with the **BTs** and **BUs**. There was also a much larger proportion of positive respondents from the 300 subjects who saw the treatments *With Goals*, 13%, than from the equal body of subjects who saw them *Without Goals*, 7.7%. Therefore the presence of *Goals* seems to markedly increase the number of positive respondents.

An analysis of variance was initially tested across the six Main Display Treatments because this was the more conservative approach to testing for differences between the resultant (number of positive minus negative) positive comments. As shown in the preceeding table, the analysis across the Main Display Treatments did not yield significant results below the .05 level. An analysis of variance across the Principal Display Treatments did, however, produce significant results, and a similar analysis between the treatments *With Goals* and those *Without Goals*, yielded results which were close to being significant at the .05 level. The differences between the number of resultant positive comments for the Principal Display Treatments were significant at the .038 level (DF=2, F=3.290), and the treatments *With* and *Without Goals* were significant at the .052 level (DF=1, F=3.795). However the failure to find significant results between the Main Display Treatments inevitably suggests that the interpretation of the subsequent results must be regarded with caution.

The **MTs** received the largest number of resultant favourable comments. They also received the most suggestions to space out the objects. The **BUs**

received a below average number of resultant favourable comments and the least number of complaints against the overcrowding of objects. The **BTs** received the smallest number of resultant positive comments, and the most emphasis on the need for the vertical separation of themes, as well as on the need for more space in general. The treatments *Without Goals* also had more resultant positive comments than the same treatments *With Goals*.

Overall, many visitors requested the use of additional photographs, illustrations and reconstructions; maps showing the locations where objects were found; and time lines to give a visual picture of changes. A number also complained because the "Ages" were arranged chronologically from right to left, i.e.

Late Iron Age- Early Iron Age- Bronze Age- Stone Age

instead of the way we usually read, from left to right.

6.4.1. COMMENTS- GENERAL IMPACT OF THE DISPLAY CASE DESIGN VARIATIONS

excellent- the main activities of man carried through graphically and in easy to understand fashion. <MT 17>

Spaced out more- give the material full value- better lighting. <MT 91>

It attracted my attention and was an excellent starting point. It made me think I'd buy a book on the subject. <MTG 116> (-note: she was amazed because she usually whips around such places when she goes with her husband and then sits in the corner and waits. But she found this case so simple and easy to follow that she stayed interested in it for quite awhile.)

it's good because it has everything organized into various subheadings, suitable for comparison. <BTG 54>

Basically good, but a bit too cluttered- needs to be spread out more. <BTG 36>

it was interesting and held your attention. <BU 7>

I thought the case was quite well laid out but must confess that I came in to see St. Ninian's Treasure and only glanced into this case. <BU 14>

ordinary visually but written information was quite useful <BUG 113>

6.4.1.1. Overview of Visitor Comment Groupings

In this section examples of visitors' comments are presented in order to demonstrate the range of feedback encountered from just one of the Principal Treatment Groups, in this instance, the **MTs**. The comments are broken down and grouped into relevant categories. The comment categories, listed in order of appearance, are:

- General Impact of the Display Case Design Variations
- Concepts Learned: No Warfare, Social Integration
- Relevance to Personal Experience
- Comments on Display Techniques: General, Too Crowded
- Use of Introductory Case
- Use of Questionnaire
- Suggestions for Improving This Display
- Opinions and Expectations

As well as looking at the comment groupings, it is also interesting to note the types of comments which are expressed. For example, in the "No Warfare" grouping, the first comment demonstrates "Rule Learning" (Gagne), when concepts have been absorbed, internalized, and utilized to create a 'new' hypothesis. The second comment is a rewording of facts absorbed from the display case. The third is a repetition of phrases used in the display. The fourth comment demonstrates a personal reaction.

6.4.2. CONCEPTS LEARNED

6.4.2.1. No Warfare

the chronology of wars showing no weapons of war until a surplus of goods was produced <MT 17>.

the fact that in the early Stone Age no weapons for warfare were known. <MT 48

...details such as suggestion of lack of weapons in earliest period indicates no fighting. <MT 88>

The absence of weapons in the Stone Age surprised me... <MTG 169>

6.4.2.2. Social Integration

(how they) got food- harpoon head to get fish and seals, sewing with needles, buried in urns. <MT 42>

evidence of close knit community yet each family having independence of its own. <MT 75>

the progression of tribes and influences that shaped the modern Scots. <MT 77>

the way in which crafts were made and the type of tools used. <MT 89>

weaving, beads, personal objects- jewellery, etc. <MT 91>

Burial of the baby on the swan's feather... <MTG 105>

details about the first people coming to Scotland. <MTG 108>

the fact that the Christians did not have rich grave goods. <MTG 170>

6.4.3. RELEVANCE TO PERSONAL EXPERIENCE

very interesting display- at quick glance I was comparing the periods to those I've seen in similar museums in North America (e.g. South Western early American Indian cultures). <MT 59>

we noted that the early North American Indian graves used the crouched figure. <MT 16>

the design and hafting of the bronze implements, in comparison to modern tools. <MT 59>

bronze needle, iron plough share, i.e. articles which have continued in use in refined forms. <MTG 120>

fact that method of handweaving has not changed in so much time. <MTG 121>

...the reconstruction of Jarlshof, as I remember seeing the ruins... <MTG 129>

comparing the penannular brooch in the last case with those I'd seen in the Irish exhibit in San Francisco... <MTG 161>

Silver so early on. Cremation urns- we seem to have come full circle. <MT 78>

6.4.4. COMMENTS ON DISPLAY TECHNIQUES

it did look different from the other displays in many ways: pictures, display methods, timing, etc. <MT 3>

(interested in) tools and ornaments, especially shown "in use", photographs of graves. <MT 18>

I was struck from the very first by how informative it was <MT 84>

A very good display, but questions prompt me to go look at it again. Some guidance on what to look for would be helpful (unexpected/unusual) <MT 85>.

I liked it- particularly the brief outlines. <MT 88>

I think the display is basically a very good idea, especially as it gives specific points for the visitor to concentrate on, enabling them to grasp (in theory) quickly the essentials of what the display is trying to say, though I did think it was all a little too compressed. <MTG 101>

Catches the eye when entering the room. A good display. <MTG 106>

easy to follow developments because of top labels of case. <MTG 115>

I loved the organization grouping, both as to general archaeological period and to specifics within them (food, weapons, etc.) I would like a few more examples of craft techniques, farming methods and other factors of everyday life. <MTG 118>

good informative layout showing basic evolutionary traditions of objects displayed. Examples well-chosen, good representatives of types. For size of area very good. <MTG 127>

it is a useful layout to show the divisions between the main periods. No constructive comments other than possibly brighter type needed. <MTG 141>

interesting layout as opposed to the other show cases. When there are too many of the same thing it gets boring and when you get to this stage you do not even bother reading the information cards. <MTG 166>

methods of using photographs to illustrate the artifacts as they were found was extremely helpful. -excellect arrangement. <MTG 188>

the first case is well situated for starting. I liked the vertical organization by date. After a short while I discovered the horizontal development of tools, weapons, ornaments, food which I began to follow with interest. It might be more apparent if lines were added showing this horizontal development- without destroying the artistic quality of the display. Please- more of

this kind of display. <MTG 195>

6.4.4.1. Too Crowded

The display case very good. But there is so much to look at in the overall display that I felt it might be advantageous to show perhaps fewer objects and concentrate on one particular era. <MT 23>

Bit crowded- I'd pump for greater space for one "theme" at a time- like farming or weapons or tools- and possibly show the relationship to what was happening elsewhere- like whether a development was indigenous or brought about by external events. (like Romans and ironmongery!) I guess I'm arguing for about four more display cases to finish telling the story! <MT 55>

arrangement balanced. We hurried through it. Perhaps it is too crowded. <MT 71>

quite good, but a bit crowded. Such displays could be keyed to similar arrangements of other artifacts. <MTG 112>

1) the arrangement suffered from the size and style of the presentation case.
2) also the number of items was not consistent with the size of the case.
3) the items on display were individually interesting, therefore I suggest that each unit should be given more space. And there should be a space between each of the four Ages on display... <MTG 125>

spaced out more- maybe two Ages in the case. <MTG 153>

I thought it was nicely displayed but a bit too crowded. <MTG 169>

the arrangement of this display was good, and not too cluttered. A display of this kind should be clear and simple for all different age groups to understand. I feel it would have been better if more objects had been displayed in the four Ages, and each case extended slightly to accommodate for this. <MTG 124>

it was a very good display. Not overcrowded, and easily followed. <MTG 194>

6.4.5. USE OF INTRODUCTORY CASE

I think it's a good idea to start with this case as it charts the very beginnings of Scottish history. <MT 12>

the explanations were very good but they could be expanded a little more. The visual effect is pleasing and orderly. Please expand the number of cases provided to outline the history of Scotland. <MT 77>

visitors should be directed to this display case first- If they are- it is not obvious enough. Well done! <MTG 137>

interesting- as I have two boys- ages 5 and 7 years and I was able to explain aspects of life to them over a period of years, rather than just look at a case of similar objects with no time sequence. Photographs were very popular. -more household items. <MTG 155>

liked it very well- only- I saw it after 25 minutes of viewing the room backwards and was too tired to read it carefully. I

wish I had started with it. It would have lent much more meaning to what I saw- I would have thought more in terms of "Ages". It gave a good summary of the main features of the different periods. <MTG 171>

6.4.6. USE OF QUESTIONNAIRE

I think that this questionnaire is an excellent idea and a lot of fun. It is a good guide for visitors and a way to learn, understand, and retain knowledge. It was interesting and helpful- should be available to visitors and "browsers" upon their seeing the museum. <MT 4>

Display- helpful- whetted interest- clear. I- granny- came with 7 year old grandson at his request. A tour with questionnaire for these bright children and ignorant elders would be useful to both. Very simple questions leading to regular exploratory future visits would be useful. Questions to be seen at answering of course! Thanks for this set of questions. All the best for your project. (signed) <MT 78>

(display improvement would) probably (be the) ability to obtain this type of questionnaire at end of case at all times to make you go back and really look. <MTG 175>

well displayed and informative, questionnaire very good idea. <MTG 185>

...(There should be a warning about questionnaires) WOTTA DAY <MT 88>

6.4.7. SUGGESTIONS FOR IMPROVING THIS DISPLAY

Perhaps would be interesting to have a model of a typical dweller in each age- clothes, ornaments, housing etc. Bit more

human interest to catch attention. <MT 12>

Generally O.K. Possibly fewer objects and more use of colour and bolder lettering might help in making display more attractive and more easily assimilated. <MT 13>

use of a different colour for the different aspects of life across the display would help to provide a unifying aspect to these (included a drawing of theme bands). <MT 57>

I was confused by the right to left display. Perhaps labelling it with "Start Here" would prevent others from studying it backwards as I did. <MTG 122>

...4) the lettering is very clear, though it might be an idea to introduce some variation (?)- for each Age a different letter type?

5) items which are too low may not get the same attention at a glance, as those at eye level? Some way of counteracting this could be found. (for example see the Royal Scottish Museum's display of silverware).

6) display cases in which there is no direct involvement by the viewer, in the case of younger children prove less appealing. Allowing articles which are suitably unbreakable for physical contact, produces a more involved public. (For example many zoo's have animals which children can feed, and stroke etc.) <MTG 125>

...if can obtain larger area then perhaps a bit more information to fill in gaps in time sequence. Explain Iron Age in Britain not quite same as in type sites of Mediterranean. (signed) <MTG 127>

the arrangement was good although the display case itself is not very good. The use of a projector and sound track could be used. e.g. as used in Koln Museum. <MTG 128>

clearly laid out- perhaps more vivid accompanying illustrations would impress it more deeply on the memory of visitors. <MTG 173>

6.4.8. OPINIONS AND EXPECTATIONS

It's always the case that many items of display and much detail in the labels obscures the basic development- this was not so apparent in this display- but then one can not expect to teach much or to raise interest in those who wander around displays. <MT 52>

to some extent, my desire in coming was merely to see the artifacts and implements an display, as being "interesting in themselves". The relation of items to specific historical times is of more interest to the historian than the casual museum visitor, although this seems to be the object of this quiz. Development is obvious merely by seeing one case after another, once the artefacts are arranged in historical order, and any regressions show themselves too. I do not think dates matter much. <MT 69>

I am amazed at my lack of observation but then I do not think your display was to blame. I was more interested in the actual article itself rather than the information about it. Possibly another time or in another mood I'd be interested in reading about the various things but this takes time and in reality how many people want to spend hours in a museum looking at the exhibits. For my own part, and no doubt for many others, a museum visit is a "filler" rather than a "total" activity hence I feel the presentation should be very simple and direct- after all, are not guided tours provided to give the "star- studded" version. (If not, then they should be it's so much easier to communicate with the spoken as opposed to the written word.) Sorry I flunked your test re the information on the articles, but by the same token this questionnaire has made me consider why I come to such places as this. <MT 86>

...Not needing an introduction to archaeology I found I gave it only a very cursory look. Personally I think National Museums should cater more for the needs of the specialist, leaving the basics to local museums. <MTG 101>

being pressed for time did not look very long. <MTG 105>

although interested, took little in as preoccupied with feeling "ill". <MTG 158>

more informative than other displays- sequences very good- would have taken in more if I had not seen so much before! <MTG 187>

6.5. STATISTICAL COMPARISON OF TEST CONDITIONS AND VISITOR CHARACTERISTICS WITH THE DEPENDENT VARIABLE RESULTS:

Since the results of the analyses of variance between treatments for *Knowledge Gain* and *Written Comment* scores are somewhat equivocal, further tests of these variables were necessary. A multiple regression analysis was performed to determine the relative importance of selected independent variables on the results of the *Knowledge Gain* and *Written Comment* scores and, in part, to test whether the effects of the Principal Treatments' differences on these outcomes could be explained by these other independent variables. The independent variables were chosen because they had either been shown in Chapter 5 to differ significantly between treatments, or because they might be expected to have an effect on the questionnaire results. These independent variables were: Weather (sunny, cloudy, rain); Floor Conditions (quiet, noisy); Day of the Week (weekend, weekday); Residence (Edinburgh, not Edinburgh); Visit (first time, repeat visit); avowed Interest in Scottish archaeology (very to moderate interest, slight to no interest); avowed Knowledge of Scottish archaeology (very good to fair knowledge, little to no knowledge); age when full-time Education completed (0-18, 19 or over, student now). In addition, the three Principal Display Treatments, and the treatments *With Goals* and *Without Goals* were also entered into the regression analysis in order to assess their separate effect. These variables were not entered into the equation in any pre-selected order.

TABLE 6.5- Regression of Knowledge Gain and Resultant Written Comments on: the Principal Display Treatments, Goals, and Selected "Control" Variables.

	KNOWLEDGE GAIN		WRITTEN COMMENTS	
	B	F	B	F
GOALS (1=present,0=absent)	0.26	1.10	-0.25	1.39
MTS (1=MTs,0=BTs or BUs)	0.80	6.41*	0.45	2.84
BUS (1=BUs,0=BTs or MTs)	0.66	5.04*	0.11	0.18
SUN (1=yes,0=rain or clouds)	0.11	0.18	-0.25	1.23
RAIN (1=yes,0=sun or clouds)	0.17	0.18	0.11	0.10
QUIET (1=yes,0=noisy)	-0.05	0.03	0.33	1.89
WEEKEND (1=yes,0=weekday)	-0.12	0.22	0.21	0.86
EDINBURGH (1=resident,0=not)	1.00	6.83*	0.14	0.19
FIRST TIME VISIT (1=yes,0=repeat)	1.00	11.55*	0.74	8.20*
INTEREST (1=very-mod,0=slight)	1.44	31.15*	1.03	22.36*
KNOWLEDGE (1=very-fair,0=little)	0.82	8.70*	0.68	8.32*
EDUCATION END (1=19+,0=pre19)	1.79	45.59*	0.58	6.65*
STUDENT (1=yes,0=no)	2.20	37.06*	0.71	5.39*
(CONSTANT)	0.92		-0.60	
R SQUARE=	0.25		0.14	
OVERALL F=	15.36		7.39	
NUMBER=	600		600	

* =indicates that the significance level was less than 0.05.

Table 6.5 reports the results of the analysis. Similar analyses of the *Initial Attraction* and *Interest Retention* variables, however, couldn't be conducted due to the fact that these two variables were based on a larger sample than the 600 for whom the relevant control variables were measured, and therefore were not calculated separately for each subject, but were instead a reflection of the cumulative reactions of all visitors to the gallery within each treatment.

The preceeding table suggests that the presence or absence of *Goals* did not have a significant effect on either *Knowledge Gain* or *Written Comment* scores. The **MTs** and **BUs** treatments did however prove to have a significant

effect on increasing *Knowledge Gain* scores, even with the other independent variables included in the analysis. This result is consistent with the findings from the analysis of variance reported in table 6.3.

None of the Test Conditions (Weather, Floor Conditions, or Day of the Week testing took place) proved to be significant in the regression run. (Throughout this thesis, the term "significant" refers to significance levels which were below 0.05.) Three of the Visitor Characteristics were however highly significant for both *Knowledge Gain* and *Written Comment* scores. These were: avowed Interest in Scottish archaeology, Knowledge of Scottish archaeology, and age when full-time Education was completed. Therefore visitors who expressed a moderate to great Interest in Scottish archaeology, or a fair to very good Knowledge of that subject, or who had either completed their Education after the age of 18 or were still students, were likely to obtain higher *Knowledge Gain* and *Written Comment* scores than were their counterparts. As shown in Chapter 5, Interest, Knowledge, and Education levels did not differ significantly between the treatment groups. Table 6.5 confirms that the significant differences between the Principal Treatment groups cannot be explained by these variables, nor can they be explained by the presence of First Time visitors and those who were Resident in Edinburgh, both of which also had significantly increased *Knowledge Gain* scores.

It is also interesting to conduct a correlation analysis on three types of visitor behaviour patterns which have been suggested by some museum evaluators (Betchel, 1967; Shettel, 1968; Munyer, 1969; Screven, 1974) to have an impact on learning. These behavioural visitor characteristics may have been influenced by the treatments being viewed, and for this reason have not been included in the previous regression analysis. They were: viewing sequence (right to left was the correct chronological sequence), i.e. the sequence in which the four panels of the chronological display were seen (right to left and

repeat, right to left, left to right, random); viewing time, the length of time that the tested subjects viewed the display case (from 45 seconds to 25 minutes); and interactions, the type of interactions that visitors engaged in when looking at the display case (much talk-notes-gestures, talk about a specific topic, little talk). In addition, the direction in which visitors moved immediately after they entered the gallery (to the test case, left, straight, right) is also discussed in this section, since it too was shown to reflect visitor behaviour modifications, due to the influence of the display treatments on show (it was significant at the 0.0052 level). Two other visitor behaviour patterns were not analysed in depth, in the subsequent table, because they did not demonstrate significant differences between treatment groups. These visitor behaviour factors were: goaltime, with a sigificance level of 0.1861, and the direction of approach to the experimental display case (from the right or the left), with a significance level of 0.7500.

TABLE 6.6- Correlation of Visitor Viewing Sequence, Viewing Time, Interactions, and Initial Traffic Direction, with Knowledge Gain and Written Comments:

	KNOWLEDGE GAIN:			WRITTEN COMMENTS:		
	MTs	BTs	BUs	MTs	BTs	BUs
Viewing Seq.	-0.12*	-0.06	-0.15**	-0.19**	-0.09	0.01
Viewing Time	0.22***	0.36***	0.34***	0.08	0.11	0.33***
Interactions	0.07	0.16	-0.14	-0.27*	0.12	-0.46***
Direction	0.08	0.06	0.17**	0.16	-0.06	0.07

* - indicates that the significance level was less than 0.1.
** - indicates a less than 0.05 significance level.
*** - indicates a less than 0.01 significance level.

Table 6.6 shows partial correlations within each Principal Treatment group. This type of analysis is necessary because the effects of different visitor behaviour patterns might vary across treatment groups. The correlation

coefficient analysis showed a strong positive relationship between viewing time and *Knowledge Gain*, across the three Principal Display Treatment groups, and between viewing time and *Written Comment* responses in the **BUs** treatment group. Chi square analysis of viewing time with *Knowledge Gain*, revealed a significance level of 0.0000 across the Principal Display Treatments, and, within the treatments, of 0.1010 for the **MTs**, 0.0216 for the **BTs**, and 0.0050 for the **BUs**. Therefore there appears to be a stronger relationship between long viewing times and high test scores in the **BUs** treatment, than in the two other Principal Display Treatment groups. These results are consistent with the suggestion that visitors needed more time to absorb information from **BUs** than from the other treatments. **BUs** only differed from the other treatments in its lack of bold headings, separated paragraphs, and a structured object arrangement. It may be suggested, therefore, that bold headings, separated paragraphs, and a structured object arrangement facilitate information retrieval from a display case. These implications suggest that observing visitor viewing times, as a substitute for visitor testing, (as was tentatively proposed in Chapter 4) might mask potential learning difficulties.

A chronologically correct viewing sequence appeared to have a slight influence on increasing visitor test scores in the **MTs** treatment, and a significant impact in the **BUs** treatment. It may be that when when there is little immediately obvious structure in a display, as in the **BUs** treatment, a correct viewing sequence is of more importance in generating high test scores, than it is in overtly organized displays. The presence of *Goals* also slightly increased the number of subjects who viewed the **BUs** displays in the correct sequence, but had no significant impact upon the other two Principal Treatment groups. One may infer from these results, that, in the absence of other major exhibit organizers within a display, the presence of *Goals* may have some influence in reinforcing the use of a correct viewing sequence. The correct

viewing sequence also was significantly correlated with a high number of positive comments in the **MTs** treatment.

There was a slight correlation between a high number of visitor interactions and a high number of positive *Written Comments* in the **MTs** treatment, and a significant correlation in the **BUs** treatment. The number of interactions appeared to have no significant impact on *Knowledge Gain* scores. Visitor interactions of any form only occured in approximately one quarter of the sample in each of the Principal Treatment groups.

The initial direction of visitor movements after entering the gallery appeared to have no significant effect on the results of the questionnaire, in all but one treatment group. The **BUs** treatment showed a significant relationship between visitors who initially travelled towards the right hand side of the museum, away from the experimental display case, and high *Knowledge Gain* scores. However the proportion of visitors who initially turned to the right was very small, (two-thirds went directly to the experimental display case) and was not of a sufficient size to use as a basis for significant inferences.

6.6. CONTROL GROUP EVALUATION MEASURES:

1.1. Initial Attraction of Control Groups-

TABLE 6.7- Initial Attraction- Results of the Main Display Treatment means and Control Groups:

DISPLAY TREATMENTS	% OF INITIAL ATTRACTION	TOTAL NO.OF VISITORS	NO.OF INITIAL VIEWERS
Main Display Treatments:			
MEAN	32.5%	452	146
Control Groups:			
MTa	30.5%	774	236
RMT	29.0%	641	186
YELLOW CONTROL	11.4%	472	54
GREEN PILOT CASE	32.5%	237	77

The ***MTa*** Treatment, with its small **MT** headings, and the ***RMT*** Treatment, which duplicated the **MT** format at the end of this study, had mean *Initial*

Attraction values which were close to those of the Main Display Treatments. The ***Green Pilot Case***, with its prominent small labels, was close to the group mean for the other green display treatments, the Main Display Treatments. It's *Initial Attraction* value was 32.5%. However, when a green sign was added to the ***Green Pilot Case***, the mean fell to 22.9%! The ***Yellow Case Control*** displays, as a group, had a significantly lower mean of 11.4%. The original ***Yellow*** display case, which had no sign beside it, was far lower still, with a mean of 8.9% for its *Initial Attraction* value. However, when a buff coloured sign was added beside the case, the mean rose to 23.8%, putting it on a par with the ***Green Pilot Case*** and its sign.

2. Interest Retention of Control Groups–

TABLE 6.8– Interest Retention time– Results of the Main Display Treatment means and Control Groups:

DISPLAY TREATMENTS	MEAN TOTAL INTEREST TIME	TOTAL TIME	TOTAL NUMBER OF VIEWERS
Main Display Treatments:			
MEAN	113 seconds	169060"	1501
Control Groups:			
MTa	70 seconds	30957"	445
RMT	102 seconds	31810"	313
YELLOW CONTROL	25 seconds	5054"	198
GREEN PILOT CASE	43 seconds	3878"	90

The Principal test Treatments (**MT–BUG**) scored significantly higher, both individually and as a group, than did any of the Control Treatments, apart from **RMT**. The *Interest Retention* times denoted a marked difference between the different types of Control Group displays. The original ***Yellow Case Control*** had

a mean total interest time of only 25 seconds while the ***Green Pilot Case***, with its numerous small white labels, had a mean time of 43 seconds. The ***MTa*** group (which was identical to **MT** apart from the use of smaller label category headings) spent an average of 70 seconds looking at that display treatment. Therefore it may be that the use of numerous small identity labels, which contrast with the background display colour, and small headings, will detract from the time visitors spend looking at a display format. A better display format would appear to be one which has homogeneous identity label and background colours, and large headings.

3. **Knowledge Gain of Control Groups**–

TABLE 6.9– Knowledge Gain scores– Results of Main Display Treatment means and Control Groups:

DISPLAY TREATMENTS	KNOWLEDGE GAIN MEAN TOTAL TEST SCORES[max= 18]	HIGHEST SCORE [max=18]	NO. OF NULL SCORES	NUMBER OF VISITORS
Main Display Treatments:				
MEAN	4.77 (26.5%)*	14	11	100
RANGE	1.19	8	10	600
Control Groups:				
MTa	5.37 (29.8%)	15	6	100
RMT	4.90 (27.2%)	15	10	100
PILOT(unseen)	3.81 (21.2%)	11 (1.7%)#	2 (3.4%)#	59
PRE-TEST(unseen)	2.73 (15.2%)	6 (9.1%)	3(13.6%)	22
POST-TEST	7.09 (39.4%)	16 (4.5%)	1 (4.5%)	(same 22)
ASK STUDY	11.18 (62.1%)	18 (5.3%)	0	76
TEST SELF	10.35 (57.5%)	18 (15%)	0	20

Designations for the entire column:
* =the percentage correct out of the maximum possible test score.
# =the percentage of subjects in this group which attained this score.

Much effort during the pilot-testing phase went into re-wording the questions and arranging the choices to minimize the number of correct guesses which took place. Various Pilot Questionnaire formats were tried out on visitors. The final questionnaire version was used for all groups, apart from the ***Pilot Questionnaire (Unseen)*** and ***Alleval Rating*** Control Groups, although all of the test scores reported in the ***Pilot Questionnaire (Unseen)*** group used a questionnaire format which was similar to the final version.

The only Control Group which used the final questionnaire format, but did not allow the visitors to see the display case, was the ***Pre-Test*** Group. Therefore this group is important as a basic representation of the number of correct responses which could be attained through correct guesswork and pre-display treatment knowledge.

It was not possible however, given the space restraints imposed upon question format by the decision to limit the questionnaire to two pages in length, to reduce the average number of correct responses to the Control ***Pilot Questionnaire (display unseen)*** and ***Pre-Test*** below 16%, without increasing the complexity and narrowing the focus of the questions themselves. The mean test score of 2.73 on the Control ***Pre-Test*** would seem to imply that approximately one-sixth of the questions could be answered by the average visitor to that museum who had never seen the display case.

The Control Groups which did not see the display case had predictably low mean Test Scores. The ***Pre-Test*** group had a mean Test Score of only 15.2% correct, and a High Score of 6 points. The ***Pilot Questionnaire*** test group, who also never saw the display case, had a low mean Test Score of 21.2% correct. On the other hand, predictably high mean Test Scores of 62.1% correct were attained by the ***Ask Study*** group, who were asked to study the display, knowing they would be tested on it. In this group 5.3% attained the maximum High

Score and none had a Null Score. The group who completed the test while looking at the display, the ***Test (Your) Self*** group, had high Test Scores of 57.5% correct and the largest percentage of maximum High Scores, 15%. The ***Post-Test*** demonstrated a 24.2% increase over the ***Pre-Test*** score and a huge jump as well in the highest score attained, which went from 6 to 16 points. The ***MTa*** and ***RMT*** Treatments had mean Test Scores, High Scores, and Null Scores close to those of the **MTs**.

The significance of the Control Group Test Scores are as follows: The low ***Pre-Test*** scores demonstrated that the average casual visitor to the museum was only likely to already know, or be able to guess, 2.73 of the test questions correctly. Therefore the test did not deal with common knowledge or easily guessed questions. The ***Ask Study*** and ***Test Self*** groups showed that all of the answers could be found in the display case and suggests the mean scores which might be attained by very interested and motivated casual visitors. The mean Main Display Treatment score falls in between these parameters, and closer to the group who did not see the display. This may suggest that visitors do not often read and observe a display case thoroughly, as indeed the viewing times indicate, but that a well designed case will help them to increase their knowledge and understanding of the areas which do interest them.

4. Written Comments of Control Respondents–

TABLE 6.10– Number of positive minus negative respondents and comments for Main Display Treatment means and Control Groups:

DISPLAY TREATMENTS	RESULTANT RESPONDENTS	RESULTANT COMMENTS
Main Display Treatments:		
MEAN	10	143
Control Groups:		
MTa	27 positive	196 positive
RMT	14 positive	140 positive

The subjects in the Control Groups which did not see any of the display treatment variations, i.e. the ***Pilot Questionnaire (Unseen)*** and ***Pre–Test*** Groups, did not, of course, provide written comments about a display. Three of the other Control Group subjects, those in the ***Post–Test, Ask Study,*** and ***Test Self*** groups, wrote down very few comments. It therefore remains to compare the Main Display Treatment mean with those of the vociferous subjects in the ***MTa*** and ***RMT*** treatment groups. The re–test of the ***MT*** treatment at the end of the study, ***RMT,*** had significantly more positive respondents (14) than did **BU** with 5, although not quite as many as **MT**, with 21. This would suggest that the great difference between the means for the **MT** and **BU** treatments is not due to seasonal changes in the visiting population.

5. Comparison of MTa, RMT, and MT

The Main Display Treatment **MT** was tested at the very beginning of the main study. It was also tested at the end of the main study, when it was called ***RMT,*** as one way of uncovering any major seasonal differences in visitor actions and characteristics. The same treatment was furthermore slightly

modified to contain smaller headings, and tested prior to the main study, as ***MTa***. A comparison of these last two treatment results may therefore suggest whether or not bold headings produce a quantifiable increase in communication effectiveness.

RMT had an *Initial Attraction* value of 29%, which was substantially lower than **MT**'s 35.7%. It also had a lower *Interest Retention* time of 102 seconds, compared with **MT**'s 114 seconds. ***RMT***s mean *Knowledge Gain* score of 4.90 points was somewhat lower than **MT**'s 5.07 points, but ***RMT*** had a higher High Score of 15 (versus 13) and a comparable Null Score of 10. There were 14 resultant positive respondents and 140 resultant positive comments for ***RMT***. These were lower, on both counts, than the 21 resultant positive respondents and the 194 resultant positive comments for **MT**.

Although the ***RMT*** results were somewhat lower than the **MT** results on most counts, they were still generally on a par with, or above, the non-MT treatments. Its *Initial Attraction* value was similar to that of **MTG** and all of the other treatments apart from **MT** and **BUG**. Its **Interest Retention** value was lower than the **MTs** and **BUs**, but higher than the **BTs**. Its *Knowledge Gain* score was higher than all of the non-MTs apart from **BUG**. ***RMT***s Written Comment results were substantially higher than all of the non-MTs. One might postulate then that *Interest Retention* scores may be the most vunerable to seasonal visitor changes. Alternatively, it could be that ***RMT*** represents a different visitor population from those in the rest of the study, since it was tested at the start of the school summer holidays.

MTas *Initial Attraction* value of 30.5% and mean *Interest Retention* value of only 70 seconds were also substantially below **MT**'s equivalent of 35.7% and 114 seconds. However it experienced a relative increase in its questionnaire results, from **MT**'s 5.07 mean points test score and 21 and 194 resultant

positive respondents and comments respectively. ***MTa*** scored 27 points and had 27 resultant positive respondents and 196 resultant positive comments. It may be that fewer people in the general population were interested in looking at ***MTa*** for a long enough period of time to qualify to be tested, so consequently, those who did complete the questionnaire were more interested and motivated than were the more general subjects who were tested under **MT**. These results and this conclusion suggest that bold headings did substantially increase the amount of time visitors looked at the display treatment.

6.7. SUMMARY

The Principal Display Treatments, **MT–BUG**, had reasonably similar *Initial Attraction* values. As a whole, they attracted a third of the visitors as soon as they walked into the hall.

On average, those of the one and a half thousand people who visited the hall during the Principal Treatment experiments, and who looked at the display case, did so for approximately two minutes. The **BUs** and the **MTs** display treatments had substantially longer mean *Interest Retention* times than did the **BTs**, or any of the Control Groups, apart from ***RMT***. In addition, the mean *Interest Retention* time was also greater for those display treatments *With Goals* than for those *Without Goals*.

The subjects' *Knowledge Gain*, as represented by their total test scores, allowed for a maximum of 18 points. However the *Knowledge Gain* values for the six Main Display Treatments all clustered around the 5 point mark. Overall, both the **MTs** and the *With Goals* Treatments had slightly higher Test Scores, High Scores, and Null Scores than did the other Main Display Treatment combinations.

The *Written Comment* responses, however, showed very definite differences

in the subjects' reactions to the display treatments. After the number of negative respondents had been subtracted, the **MTs** were left with 46 resultant positive respondents, the **BUs** with 13, and the **BTs** with only 3. Expressed as percentages, the **MTs** had 73% more positive respondents than the **BUs**, who had 22% more positive respondents than the **BTs**. The *With Goals* Treatments, with 39 positive respondents, demonstrated a marked increase over the 23 positive respondents in the *Without Goals* Treatments. However the treatments *Without Goals* had 37% more resultant positive comments than the treatments *With Goals*. It would seem therefore that the subjects who saw the treatments *Without Goals* produced a greater quantity of positive comments, relative to the number of negative comments they produced, but that they contained a smaller group of respondents who produced totally positive comments, relative to the number of totally negative respondents, than did the *With Goals* group.

When an analysis of variance was conducted in a rigorously conservative manner across the six Main Display Treatments, it did not produce significant results for either *Knowledge Gain* or the number of resultant positive *Written Comments*. However subsequent analyses of variance across the three Principal Display Treatments, and between the *With Goal* and *Without Goal* treatment groups, did suggest significant trends. The differences in both *Knowledge Gain* and in the resultant number of positive *Written Comments* were significant below the .05 level for the Principal Display Treatments. Furthermore, although the presence or absence of *Goals* did not significantly affect *Knowledge Gain* scores, they were very close to being significant for the number of resultant positive *Written Comments*. In view of the failure to find significant results across the Main Display Treatments, it is accepted that the subsequent results must necessarily be equivocal, and subsequent inferences must be cautious. However the direction of difference is consistent with the expectations of the study that significant differences would exist between the

Principal Treatment groups and between the presence or absence of *Goals*. Therefore despite the necessarily equivocal nature of these subsequent results, they may still be viewed with some measure of confidence since they express the same direction of difference across the Principal Display Treatments, and to a lesser extent, across treatment groups *With* and *Without Goals*.

CHAPTER 7

DISCUSSION

7.1. BRIEF SUMMARY OF RESEARCH METHODS AND RESULTS:

The principal aim of this study was to evaluate the effectiveness of opposing object, label, and goal presentations to communicate with, that is, to attract, interest, and instruct, the casual museum visitor. The contrasting object formats were: an unstructured display versus a thematically grouped one. The text variations were: single Block labels versus a series of multiple labels with bold headings. Goal statements were either present or absent from the other format combinations. The selected treatment combinations were successively varied within one museum display case, and were evaluated using observation of visitor movement patterns, timed viewing, written test scores, and visitor comment responses.

Although the test conditions varied between display treatments, there is little indication that these variations had any major bearing on the evaluation measurement results. Additionally, most of the subjects' demographic characteristics were reasonably similar, and again, there was no indication that the differences which did exist unduly influenced the evaluation results.

The evaluation measurement results were reasonably conclusive. The *Initial Attraction* values were similar for all of the Principal Display Treatments, as were the *Knowledge Gain* scores. The *Interest Retention* and *Written Comment* responses, however, did vary significantly. The **MTs** Treatments, with their stratified label and object formats, attained a high *Interest Retention* time, and a substantially higher number of favourable comment respondents than did any of the other treatment groups. The **BUs** Treatments, with their single main

label and random object arrangement, attained the highest *Interest Retention* time. The **BTs** Treatments, with their single main label and stratified object arrangement, received the lowest *Interest Retention* and *Written Comment* scores. There were also significant improvements in the *Interest Retention* and *Written Comment* responses when the treatments *With Goals* were displayed.

8.1.1. DETAILED RESULTS OF EVALUATIVE MEASURES:

Four measures were used to evaluate the communication effectiveness of the alternative display treatment formats in this study. These evaluative measures directly reflect the responses of the museum visitors towards the display treatments which they saw. The proportion of visitors who had an *Initial Attraction* to the display treatment, as soon as they entered the Gallery, was recorded, as was the length of time each visitor in the Gallery looked at the display treatment, the *Interest Retention* measurement. The visitors' *Knowledge Gain* measurement reflects their scores on a test which was designed to demonstrate knowledge gained from looking at the display case treatments. Visitors' *Written Comments* provide first-hand information on the subjects' perceptions of the display treatments.

8.1.1.1. Initial Attraction

The mean *Initial Attraction* values for all of the Main Display Treatments, as well as the ***MTa, RMT*** and ***Pilot Cases***, were remarkably similar. Apart from two aberrations, they all attracted close to 30% of the visitors to the Gallery as soon as they walked in through the door. The two deviants were **BUG**'s 38% and **MT**'s 36%. These two display treatments were also exceptional in having by far the largest numbers of both Edinburgh residents and repeat visitors in their sample groups. It may be that the high *Intial Attraction* proportions in these two display treatment groups could be accounted for by local residents, who were already familiar with the Gallery, and who were attracted to a new

display arrangement.

The advantages and disadvantages of a sign, directing visitors to look at the display case, have an obvious bearing on *Initial Attraction* values. The Original ***Yellow*** display treatment attracted a much greater proportion of visitors after a buff-coloured sign was placed beside it. Its mean *Initial Attraction* value rose from 9% up to 24% when the sign was added. However, when a Pilot Sign, which just said "Introductory Case" was added to the Green ***Pilot Cases***, their mean *Initial Attraction* value fell from 33% down to 23%! It may be suggested, therefore, that whereas an indifferent sign may benefit a display treatment which doesn't initially appear to be interesting, or is poorly positioned, it may also detract from a display treatment which is better at attracting initial visitors.

8.1.1.2. Interest Retention

The mean *Interest Retention* times highlight three major results: 1) The internally consistent **MT** and **BU** Treatments kept the visitors' attention much longer than did the internally inconsistent **BT** display. 2) The presence of *Goals* significantly increased the length of time that visitors looked at the display treatments. 3) The use of bold headings doubled the length of time that visitors spent viewing the same display treatment with small headings (**MT** versus ***MTa***). Additionally, as might have been anticipated, the Main Display Treatments demonstrated that they could capture the visitors' attention for much longer periods of time than the could any of the Control Groups, excepting ***RMT***.

8.1.1.3. Knowledge Gain

Although there was little demonstrable difference between the *Knowledge Gain* scores for the Principal Display Treatments, the **MTs** in particular, and the *Goal* Treatments in general, showed a slight edge over their rivals' mean Test

Scores, High Scores, and Null Scores. Predictably, the Control Groups which did not see the display case had the lowest Test Scores and High Scores (the ***Pre-Test*** and ***Pilot Questionnaire*** Groups). Also, the Control Groups which were asked to study a display treatment before they were tested on it, and those which were able to look at a display treatment while they were completing the test, had the highest mean Test Scores and High Scores, and no Null Scores (the ***Ask to Study*** and ***Test Yourself*** Groups). The results from these two groups prove that the test was constructed so that the number of correct answers which the general museum visitors were likely to know or guess before seeing the display treatments was small, and also, that they could find all the answers in the displays.

8.1.1.4. Written Comments

The "resultant" positive respondents appeared after the visitors who had included negative comments about the display treatments had been subtracted from the tested subjects who had included positive comments. The **MTs** stand apart in having a relatively high proportion of positive resultant respondents, and the **BTs** in having a relatively low proportion, when compared with the other Principal Display Treatments. The tested subjects who saw the **MTs** Display Treatments were highly attentive, and outstandingly favourable in their comments. Those who viewed the Display Treatments *With Goals* were also more positive in their comments about them, and retained their interest in them longer, than did those who were not provided with goal statements.

8.2. CONCLUSIONS:

The use of a sign changed visitor traffic patterns and directed a substantial proportion of visitors to look at the display case indicated. The presence of goals, bold headings and short paragraphs, and internally consistent displays increased the length of time that visitors remained interested in the display as

well as their positive comments about it.

The visitors' avowed preferences were for stratified object and label arrangements, and for the presence of *Goals* and other orientating statements. Taken a bit further, this conclusion suggests that visitors prefer an overtly organized display, with pointers to help them find and understand, objects and information.

The results of the study were designed to demonstrate to museum staff the effects of these differing display techniques on the specific actions and reactions of a sample of museum visitors. It is hoped that such "demonstrable" results will encourage museum staff to utilize some of these techniques to increase the communication effectiveness of their future displays. At the very least it is hoped that these experiments will reinforce the desirability of using displays to communicate rather than merely to show off objects.

It is difficult, due to the lack of comparative museum work, to reliably generalize about the applicability of the conclusions from any one study. Differences in visitor types, and their reactions to different spatial relationships within gallery areas, as well as differences in types and sizes of museums, may have a bearing on the applicability of these results. However, since good museum display policies should be dictated by museum content, it would seem appropriate to suggest that the results from this study in the National Museum of Antiquities of Scotland (or the National Museum of Scotland as it is now known) may be applied to the display policies of other museums which have a strong archaeology, social history, or even natural history content.

For the future, standardizing specific aspects of evaluation procedures and techniques to enable direct comparisons to be made between different types of museums, would greatly facilitate the advancement of knowledge within the

field of museum evaluation. Far too many studies are being undertaken in "comparative" isolation. In order to implement such strides forward, current and future evaluators need to address their attention to the following areas:

- Standardized forms could be used for recording basic observational data. Supplementary sections or forms could then be added to it to fulfill the requirements of a particular study.

- Simple conventions could also be followed for coding such things as visitor demographic and prior knowledge and interest characteristics.

- More sensitive testing procedures should be developed, both written and non-written, and compared with assessments based on rating sheets, observational studies, and interviews. In this way different assessment techniques may be evaluated for the overlapping information which they provide as well as for the uniqueness of their results.

8.3. RESULTS OF HYPOTHESES:

In order to present the focus of this study in operational terms, the following questions were posed in the Introduction. They will now be discussed individually in the light of the research findings.

- Q: **Labels**- Does the length of text and the use of bold headings attract visitors or increase their attention spans, or their ability to assimilate information?

- A: *Yes, bold headings attract visitors and increase their viewint times (based on a comparison of the **MTa** Control group which had small headings, and the bold headings in **MT**) and with a lesser degree of certainty, produce slight increases in their knowledge gain, as measured by this test, when coupled with a compatible object arrangement, as in the **MTs** treatment.*

- Q: **Objects**- If the material in a display case is organized appropriately, will many visitors attempt to carry through a train of thought, i.e. follow themes, from one section of the display case to another?

- A: *Yes.*

- Q: **Goals**- Do orientating statements (questions, goals, aims) increase visitor viewing times, knowledge gain, and the number of favourable comments?

- A: *Yes, however the slight increases in knowledge gain were not significant, as measured by this test.*

Much incidental information was gleaned concerning visitor actions, reactions, and characteristics. Such information, if utilized constructively, should help museum staff to pitch their display content to the most appropriate levels, create more attractive and informative displays, and possibly provide appropriate orientation devices to maximize the visitor's learning potential. Several of the generally important premises under scrutiny were:

- Is it possible that the person just off the street will be able to learn anything at all while standing up in a potentially noisy, crowded, unfamiliar environment?

- *Yes, see discussion of Knowledge Gain scores.*

- What length of time can visitors realistically be expected to view a display case?

- *2 minutes on average, for an important display case.*

- What is the impact of viewing sequence on maintaining interest and facilitating comprehension?

- *Random viewing shortens both.*

- Will visitors utilize directional devices such as signs?

- *Yes, if it is a suitable sign.*

1. Implications for Museum Studies

Many previous studies of learning in museums have suggested that the average visitor spends very little time looking at individual displays and reads very few labels (see Parsons, 1968; Shettel, 1968; Eason and Linn, 1975). The assumption made was that consequently little, if any, learning took place. This study challenges such assumptions in several ways, on the basis of the results of this research, on the selection of the tested samples in the afore-mentioned studies, and on the types of test intruments used.

This study found that visitors did read labels, as evidenced by their written comments and test scores. When given the opportunity of recording their preferences on an exhibit rating sheet, the majority actually asked for more label information. Their viewing times also indicated that most spent sufficient time looking at the display case for them to learn from it. The average viewing time for all visitors who looked or glanced at the display was over two minutes. For those who saw the preffered treatments (MTs and BUs), almost 9% looked at those displays for more than 5 minutes. The significant differences between the mean *Pre-Test* and Principal Display Treatment test scores also showed that demonstrable learning took place with the majority of visitors tested.

Various comments from visitors showed that they expected to learn from the display, and wanted the appropriate tools to aid their learning. The fact that visitors chose to complete the *Test Self* questionnaires also indicates that, not only were they interested in seeing what the display had to offer for them, but were also willing to use the tools at hand to assess their own learning from it. They must have viewed the learning process as part of the recreation of a museum visit, in other words as fun, since many of them took the time to search for the correct answers, with no prompting or request to do so.

One reason that the results of this study may conflict with previous research findings, is that many of the previous studies of exhibit effectiveness, particularly those which relied heavily upon written cognitive testing as a basis for assessment, failed to show that demonstrable learning occurred. There are two plausible explanations for this lack of assessed knowledge gain. One is the lack of appropriate test instruments and the other is the lack of appropriate sample pre-selection criteria.

A closer look at visitor "types" may help to explain the need for reasonable

pre-selection criteria. Many evaluators, such as Wolf and Tymitz (1978), have suggested names for visitor types which reflect their locomotor behaviour patterns when viewing exhibition halls. Wolf and Tymitz have also postulated that visitors' reactions to exhibits should be separately assessed according to the type of visitor who is reacting. For example, the responses and reactions of a "Very Interested Person" should be judged by different criteria than those used to judge a "Commuter", enroute to another area. Since Wolf's findings indicate that most visitors to a large museums are either "Commuters" or aimlessly wandering "Nomads", one would not expect the bulk of museum visitors to look at any one general exhibit long enough to markedly increase their scores on a specialist written test. Therefore it would seem sensible to establish pre-selection criteria, such as reasonable minimum viewing times over the entire exhibit area (as was done in this study), at the very least, or more rigorously, a classification of visitors based upon their traffic patterns prior to assessment.

Other studies, such as those conducted by Eason and Linn (1975) and Minda Borun (1977) suggest that alternatives, and supplements, to paper tests which require written responses, are much more effective at uncovering real knowledge gain from an exhibit. Many studies using traditional written tests, only demonstrated a 10% knowledge gain after exposure to the tested exhibit. Borun's study, which used visual stimulus and a participatory response mechasism, instead of a written test, produced an 18% increase in correct scores. Eason and Linn also found that whereas only 11% of their subjects could correctly answer a written question, 60% correct scores were achieved when the questions were presented as diagrams. In the latter example, understanding of the question was not based upon reading ability. Therefore, it may be that the use of paper and pencil tests themselves is a problem to be overcome in the refinement of assessment proceedures.

8.4. IMPLICATIONS FOR MUSEUM DESIGN:

The results of this study suggest that museum exhibitions would benefit from the use of overt exhibit organizers and goals for the visitor to follow. Goal-orientation devices increase the length of visitor viewing times and the number of positive comments, as well as slightly increasing written test scores. The use of bold headings was also shown to markedly increase visitor viewing times (see the discussion of the Control Treatment, **MTa**, with the Main Display Treatment, **MT**, results in Chapter 6). Such headings perform two main functions, that of concise exhibit organizers, providing an overview of the exhibit content, and that of repetition, encapsulating the main points to be found within the display. Therefore one may presume that repeating main summaries of exhibit content, as well as guidelines to exhibit organization, would be beneficial, whether or not they took the form of bold headings. Certainly discussions with visitors in informal interview situations during this study, indicated the need for a repetition of exhibit organizers. If the chronological ages and dates had been repeated at the bottom as well as at the top of that particular exhibit case, those visitors who were only attracted to objects at the bottom of the display case would have had a greater chance of absorbing the overall context. Certainly when given a choice, as in the ***Alleval Rating*** study, visitors chose the bold heading and short paragraph format. Signs were also shown to be effective in directing visitors' attention.

This display case was designed to allow the easy observation of visitor viewing sequences to take place. The four panels of the case were arranged in chronological sequence from right to left, rather than from left to right as in a normal reading sequence. Despite the awkwardness of following a right to left sequence, the vast majority of the tested sample, over 80%, did so, abet with many irregularities within that sequence. Roughly equal numbers of the

remaining visitors viewed the display case in a backwards sequence, from left to right, or in a random manner. Such observations reinforce the conclusion that visitors naturally follow a sequential path, at least within and exhibit case, and therefore an overtly structured display, with frequently repeated cues, has a reasonable chance of communicating sequential information to the interested museum visitor.

A further spin-off from this research is the proposition that whenever possible, an exhibit case or exhibition hall containing numerous primary text labels, which should be read in a particular sequence, should be organized from left to right, and the visitors directed to follow that path. A number of visitors in this study complained that the information was organized in the opposite sequence.

It might also be suggested that visitor traffic patterns within long halls, such as those reported by Wolf and Tymitz (1978), where viewing was linear and visitors never doubled back on their path, are different from those in single entrance/exit galleries such as the one reported in this study. In this gallery, many visitors doubled back when looking at the experimental display case and some even returned to look at the display after seeing other exhibits in the gallery.

A few museum educators such as Duncan Cameron (*Curator*, 1968) have postulated that museum exhibits do not communicate with the visitor on a linear level but rather as a pattern of discrete units. He maintains that it is up to the individual to organize it into a whole by developing his capacity for pattern recognition within the context of non-sequential learning. However the results of this study indicate that exhibits can, and do, communicate on a linear level, when the display is appropriately organized and the visitor is sufficiently interested to follow through a thematic or chronological sequence. Both the

written comment and test score results, and the observation of the sequences in which visitors viewed the display, brought about such a conclusion.

BIBLIOGRAPHY:

ABBEY, D.S., and D.F. CAMERON. "The Museum Visitor: 1 Survey Design (1959); 2 Survey Results (1960); 3 Supplementary Studies (1961)." *Reports from Information Services*, Royal Ontario Museum, Toronto.

ABLER, THOMAS S. "Traffic Pattern and Exhibit Design: A Study of Learning in the Museum." Pages 104–141 in *The Museum Visitor*, edited by Stephan F. de Borhegyi and Irene Hanson. Milwaukee: Milwaukee Public Museum Publications in Museology 3, (1968).

ALT, M. B. *Visitors to the Natural History Museum: A survey of visitors to the BM(NH) carried out in September, 1976.* British Museum of Natural History, Department of Public Services (1978).

AMERICAN MUSEUM OF NATURAL HISTORY. *A Profile of Consumer Use and Evaluation.* The American Museum of Natural History, New York, New York (1977).

BETCHEL, ROBERT B. "Hodometer Research in Museums." *Museum News*, vol. 45, no. 7 (Mar. 1967), pp.23–26.

BORHEGYI, STEPHAN F. de. "Museum Exhibits: How to Plan and Evaluate Them." *Midwest Museums Quarterly*, vol. 23, no. 2 (1963), pp. 4–5.

----------. "Space Problems and Solutions." *Museum News*, vol. 42, no. 3 (Nov. 1963), pp. 18–22.

----------. "Testing an Audience Reaction to Scientific and Anthropological Museum Exhibits." Page 7 in *The Museum Visitor*, edited by Stephan F. de Borhegyi and Irene A. Hanson. Milwaukee: Milwaukee Public Museum

Publications in Museology 3, (1968).

----------. "Testing of Audience Reaction to Museum Exhibits." *Curator*, vol. 8, no. 1 (1965), pp. 86-93.

----------. "Test Your Knowledge." *Midwest Museums Quarterly*, vol. 25, no. 4 (1965), pp. 10-11.

----------. "Visual Communication in the Science Museum." *Curator*, vol. 6, no. 1 (1963), pp. 45-57.

BORUN, MINDA. "Measuring the Immeasurable: A Pilot Study of Museum Effectiveness." Washington DC: Association of Science-Technology Centers (ASTC), (1977).

----------. *Museum Effectiveness Study: A Bibliographic Review.* The Franklin Institute, Philadelphia (1975).

BROOKS, JOYCE A. M., and PHILIP E. VERNON. "A Study of Children's Intrests and Comprehension at a Science Museum." *British Journal of Psychology*, vol. 47 (1956), pp. 175-182.

CAMERON, DUNCAN F. "Effective Exhibits -- A Search for New Guidelines. The Evaluator's Viewpoint." *Museum News*, vol. 46, no. 5 (Jan. 1968), pp. 43-45.

----------. "A Viewpiont: The Museum as a Communications System and Implications for Museum Education." *Curator*, vol. 11, no. 1 (1968), pp. 33-40.

COUTTS, HERBERT. "The Antiquities Gallery of Dundee Museum: A Visitor Survey." *Museums Journal*, vol. 70, no. 4 (1971), pp. 173-174.

DANDRIDGE, FRANK. "The Value of Design in Visual Communication." *Curator*, vol. 9, no. 4 (1960), pp. 331-336.

DOUGHTY, PHILIP S. "The Public of the Ulster Museum: A Statistical Survey." *Museums Journal*, vol. 68, no. 1 (1968), pp. 19–25.

EASON, L. P. and M. C. LINN. "Evaluation of the Effectiveness of Participatory Exhibits." *Curator*, vol. 19, no. 1 (1976), pp. 45–62.

ELLIOT, P. and R. J. LOOMIS. *Studies of Visitor Behaviour in Museums and Exhibitions: An Annotated Bobliography of Sources Primarily in the English Language.* Edited by Avis Berman, Office of Museum Programs, Smithsonian Institute, (1975).

ERWIN, DAVID G. "The Belfast Public and the Ulster Museum: A Statistical Survey." *Museums Journal*, vol. 70, no. 4 (1970), pp. 175–179.

EVANS, I. M. "Specimens or People: A Question of Communication." *Museums Journal*, vol. 69, no. 3 (1969), pp. 107–109.

FESTINGER, L. and D. KATZ. *Research Methods in the Behavioural Sciences.* New York: Holt, Rinehart and Winston, (1953) Chaps. 2,3,4.

FINNEY, I. and G. O. JONES. "Tradition and Progress: the Museological Revolution." *Museums*, vol. 23, no. 2, (1971), pp. 91–94.

GAGNE, ROBERT M. *The Conditions of Learning.* London: Holt, Rinehart and Wilson, 2nd. ed. (1970)

GENERAL REGISTER OFFICE. "Register of General Social Class" Her Majesty's Stationery Office, (1970).

GOINS, ALVIN and GEORGE GRIFFENHAGEN. "The Effect of Location and a Combination of Color Lighting and Artistic Design on Exhibit Appeal." *The Museologist*, no. 67 (1958), pp. 6–10.

----------. "Psychological Studies of Museum Visitors and Exhibits at the U.S. National Museum." *The Museologist*, no. 64 (1957), pp. 1–6.

GUNNING, ROBERT. *The Technique of Clear Writing*. New York, Rev. ed. (1968).

HUDSON, K. *A Social History of Museums*. The Macmillan Press Ltd., London and Basingstoke, (1975).

KAHN, ROBERT LOUIS, and CHARLES F. CONNELL. *The Dynamics of Interviewing: Theory, Technique and Cases*. New York (1957).

KEARNS, WILLIAM E. "Studies of Visitor Behavior at the Peabody Museum of Natural History, Yale University." *Museum News*, vol. 17, no. 14 (1940), pp. 5–8.

KLEIN, R. "Who goes to Museums?", *Illustrated London News*. April, (1974) pp. 27–29.

LAKOTA, ROBERT A. *Techniques for Improving Exhibit Effectiveness*. Office of Museum Programs, Smithsonian Institution, (1976).

LINN, MARCIA C. "Uses of Evaluation in Science and Technology Centers: Informed Decision Making, Advancing Education through Science-Orientated Programs." Lawrence Hall of Science, University of California, Berkeley. Sponsored by the National Science Foundation,

Washington, DC, (May 1975).

McWILLIAMS, BRIAN and JOYCE HOPWOOD. "The Public of Norwich Castle Museum 1971-72", *Museums Journal*, vol. 72, no. 4 (March 1973) pp. 153-156.

MARCOUSE, RENEE. "Personal Response." *Museums Journal*, vol. 67, no. 3 (1967), pp. 239-242.

MASON, TIM. "The Visitors to Manchester Museum: a questionnaire survey." *Museums Journal*, vol. 73, no. 4 (March 1974) pp. 153-157.

MELTON, ARTHUR W. "Distribution of Attention in Galleries in a Museum of Science and Industry." *Museum News*, vol. 14, no. 3 (1936), pp. 5-8.

----------"Problems of Installation in Museums of Art." *Publications of the American Association of Museums New Series.* No. 14, Washington D.C.: American Association of Museums, (1935).

----------"Some Behavior Characteristics of Museum Visitors." *The Psychological Bulletin*, vol. 30 (1933), pp. 720-721.

----------"Studies of Installation at the Pennsylvania Museum of Art." *Museum News*, vol. 10, no. 15 (1933), pp. 5-8.

----------"Visitor Behavior in Museums: Some Early Research in Environmental Design." *Human Factors*, vol. 14, no. 5 (1972), pp. 393-403.

MILES, R. S. and A. F. TOUT. "Human Biology and the New Exhibition Scheme in the British Museum (Natural History)." *Curator*, vol. 21, no. 1 (1978) pp. 36-50

MOCHON, MARION JOHNSON. "Visitor Testing in the Museum." Pages 72-74 in *The Museum Visitor*, edited by Stephan F. de Borhegyi and Irene Hanson. Milwaukee: Milwaukee Public Museum Publications in Museology, 3, (1968).

MOSER, CLAUS ADOLPH and G. KALTON. *Survey Methods in Social Investigation.* 2nd Edition, Heineman Books on Sociology, London (1975).

MUNYER, EDWARD A. *A Preliminary Survey of the Hall of Physical Anthropology, Museum of Natural History, Washington, D.C.* Washington D.C.: Smithsonian Institution Summer Institute in Display (1969).

MURRAY, C. HAY. "How to Estimate a Museum's Value." *Museums Journal,* vol. 31, (1932) pp. 527–531.

NIEHOFF, ARTHUR. "Audience Reaction in the Milwaukee Public Museum: The Winter Visitors." Pages 22–31 in *The Museum Visitor,* edited by Stephan F. de Borhegyi and Irene A. Hanson. Milwaukee: Milwaukee Public Museum Publications in Museology 3, (1968).

----------. "Characteristics of the Audience Reaction in the Milwaukee Public Museum." *Midwest Museums Quarterly,* vol. 13, no. 1 (1953), pp. 19–24.

NORTH, F. J. *Museum Labels.* Handbook for Museum Curators (B3), Museums Association, London (1957).

O'HARE, MICHAEL. "The Public Use of Art– Visitor Behavior in an Art Museum". *Curator,* vol. 17, no. 4 (1974), pp. 309–320

OPPENHEIM, A. N. *Questionnaire Design and Attitude Measurement.* Heinemann Books on Sociology, general ed. Donald Gunn MacRae, London (1972).

OWEN, DAVID E. "Are National Museums in the Provinces Necessary? A Brief Survey of Manchester Visitors to London Museums." *Museums Journal,* vol. 70, no. 1 (June 1970), p. 29.

PARSONS, LEE A. "Systematic Testing of Display Techniques for an Anthropological Exhibit", *Curator*, vol. 8, no. 2 (1965).

PAYNE, DAVID ALLEN and ROBERT F. McMORRIS. *Educational and Psychological Measurement: Contributions to Theory and Practice*. 2nd. ed., Morristown, New Jersey (1975).

PORTER, MILDRED C. B. "Behavior of the Average Visitor in the Peabody Museum of Natural History, Yale University." (Number 16 in *Publications of the American Association of Museums New Series*.) Washington, D.C.: American Association of Museums, (1938).

ROBINSON, EDWARD STEVENS. "The Behavior of the Museum Visitor." (Number 5 in *Publications of the American Association of Museums New Series*.) Washington, D.C.: American Association of Museums, (1928).

ROYAL COLLEGE OF ART, RESEARCH UNIT. *2. Report on the Relative Legibility of Alternative Letter Shapes*. London (1975).

ROYAL ONTARIO MUSEUM, COMMUNICATIONS DESEIGN TEAM. *Communicating with the Museum Visitor: Guidelines for Planning*. Toronto (1976).

SCREVEN, C. G. *The Measurement and Facilitation of Learning in a Museum Environment: An Experimental Analysis*. Washington D.C.: Smithsonian Institution Press (1974).

----------. "The Programming and Evaluation of an Exhibit Learning System." Pages 129–137 in *Opportunities for Extending Museum Contributions to Pre-College Science Education*, edited by Katherine J. Goldman. Washington, D.C.: An Informal publication of the

Smithsonian Institution (1970).

SELLTIZ, CLAIRE and M. JAHODA, M. DEUTSCH, S. COOK. *Research Methods in Social Relations*. Rev. one-vol. ed., The Society for the Psychological Study of Social Issues, New York (1959).

SHEPPARD, D. "Methods for Assessing the Value of Exhibitions." *British Journal of Educational Psychology*, vol. 30 (1960), pp. 259-265.

SHETTEL, HARRIS H. *An Evaluation Model for Measuring the Impact of Overseas Exhibits*. Technical Report No. AIR-F28-6/66-FR, Contract No. NY-66-354, Washington DC: U.S. Atomic Energy Commission (1966).

----------. "Exhibits, Art Form or Educational Medium", *Museum News*. (Sept. 1973) pp.32-44

----------. and MARGARET BUTCHER, TIMOTHY S. COTTON, JUDI NORTHRUP, DORIS CLAPP SLOUGH. *Strategies for Determining Exhibit Effectiveness*. Technical Report No. AIR-E95-4/68-FR. Pittsburg: American Institutes for Research (1968).

THORNDIKE, ROBERT LADD and LORGE. *The Teacher's Word Book of 30,000 Words*. London (1952).

WEINER, GEORGE. "Why Johnny Can't Read Labels." *Curator*, vol. 6, no. 2 (1963), pp. 143-156.

WILLIAMS, LUTHER A. "Labels: Writing, Design, and Preparation." *Curator*, vol. 3, no. 1 (1960), pp. 26-42.

WINGFIELD DIGBY, PETER W. "Visitors to three London Museums: A survey of visitors to the British Museum, Science Museum and National Maritime Museum, carried out in October and November 1971 on behalf of the Department of Education and Science," London (1974), Her Majesty's Stationery Office.

WITTLIN, ALMA. "Hazards of Communication by Exhibits." *Curator*, vol. 14, no. 2 (1971) pp. 138–150.

----------. *The Museum: Its Function in Education*, London (1949), p. 236.

WOLF, ROBERT L. ET AL. *New Perspectives on Evaluating Museum Environments: An Annotated Bibliography.* Washington, DC: Smithsonian Institution (1979).

----------. and JENNY CAVE. "Don't Brush Your Teeth Anymore. Toothpaste's Got Earth in It! A Study of the Role That Objects Can Play in the Experience of Visitors to a Museum, National Museum of Natural History, Smithsonian Institution." Washington, DC: Smithsonian Institution (1983).

----------. and BARBARA L. TYMITZ. *Whatever Happened to the Giant Wombat: An Investigation of the Impact of the Ice Age Mammals and Emergence of Man Exhibit.* National Museum of Natural History, Smithsonian Institution. Washington, D.C.: Smithsonian Institution (1978).

ZETTERBERG, HANS L. *Museums and Adult Education.* London: Augustus M. Kelley (1969).

Inert Instruction

Inert Instruction
Inert Instruction
Inert Instruction
Inert Instruction
Inert Instruction
Inert Instruction

Appendices

Appendices
Appendices
Appendices
Appendices
Appendices
Appendices

by

Tyler Buck

APPENDICES
TABLE OF CONTENTS

Appendix A. Questionnaire and Observation Sheets:

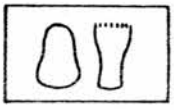
Appendix A.I. Questionnaire (reduced):

S S W — no.
M A M J___78 — var.

I need your help to find out if this type of display is useful in getting across information. Even if you are not sure of the answer, please try a guess, perhaps you picked up more than you think!

1. Please tick (✓) the one most correct time period for each item on the left.

	Stone Age	Bronze Age	Early Iron Age	Late Iron Age
(example) People first came to Scotland in the-	✓			
Early Christian Period was also the-				
First farmers arrived in the-				
Pictish artwork flourished in the-				
Cremation became popular in the -				
Warfare and Romans were here in the-				

2. 

These objects were both used for one activity-
What was that activity? _____

3. Above are illustrations of early burial types found in Scotland. Please try to put them in order from 1 to 5 - starting with 1 below the first type, and ending with 5 below the last type shown.

4. Please (circle) the two defended sites which were shown in this display case.

Skara Brae Jarlshof Dunadd Mousa

5. What did you find in this case which particularly interested you-

Please (circle) the two objects which had similar designs on them-


gold lunula


bronze dagger


penannular brooch


Kirkwall Type-A pot


jet spacer-plate necklace


halberd


bronze armlet

7. Last one! Try your skill at matching-

Please use a line to connect each item on the left with the main way you think each one changed between the Late Stone Age and the Late Iron Age-

(note- each item on left should be paired with a different one on right)

Weapons-	-far more varied in Late Iron Age
Burials-	-change in materials but not in basic shapes
Food Getting-	-not found in Late Stone Age, but often found in Late Iron Age
Ornaments-	-utilized by groups in Late Stone Age, and by individuals in Late Iron Age
Tools-	-no major change

Please answer some questions about yourself so that we can find out if a representative sample of visitors are answering this questionnaire:

(✓)

A. Is this the first time you have visited this museum?
( )Yes. ( )No- then please estimate the number of times you've visited this museum in the last 12 months _____

B. How interested are you in the early history and archaeology of Scotland?
( )very interested, ( )moderately, ( )slightly, ( )not interested

C. Would you describe your general knowledge of Scottish (or British) archaeology as- ( )very good, ( )good, ( )fair, ( )poor, ( )non-existent

D. Where are you from?
Edinburgh residents- please give district or street-
Others- please give country and town or state-

E. Are you- ( )Male. ( )Female.

F. What is your occupation and what exactly do you do in it-

G. What is your age now- _____

H. At what age did you stop being a full-time student- _____

I. Did you know you were going to be asked to fill-out this questionnaire?
( )No. ( )Yes- then how did you know-

Comments- What did you think of the arrangement of the display in this case?
How would you like to see it improved? - (use back if needed)

Please return questionnaire to box

Thank you very much for your help

Appendix A.II. Observation Sheet Example (reduced):

DATE: 15 April
DISPLAY TREATMENT: MT
DAILY TEST NUMBERS: 1 - 29
STARTING TIME: 10:10

WEATHER: mild
FLOOR CONDITIONS: quiet
OTHER:

ALL VISITORS AT ENTRANCE						VISITORS IN DISPLAY CASE VICINITY		
DAILY VISITOR NUMBER	SEX & ENTRY GROUP	INITIAL DIRECTION	VIEW TIME	COMMENTS, GOALTIME, INTERACT	TEST #	VIEW TIME	SEX & GROUP	DIRECTION AT CASE & COMMENTS
1	M }	SB			2 →	1.40	M	BA, SA ow
2	F }	SB				✓	F	
3	M }	>						
4	F }	>				.05	M	
5	Cm }	>						
6	M	★	1.25	interviewed by me after completing questionnaire - seat valid, asked the question to get correct	← 1			
7	F }	< S						
8	M }	SB				.50	M }	LIA discuss
9	F }	SB				.50	F }	
10	M }	>				✓	F	LIA
11	F }	>						
12	M }	SB				✓	M }	
13	Cm }	SB				✓	F }	
14	C	SB						
15	M }	> < ✓ SB			6 →	2.10	M }	SA ow back to look again
16	F }	> <			5 →	3.05	F }	SA
17	M }	> < ✓ SB						
18	M }	★ EIA, SA+	.35			✓	M	
19	F }	★ BA, SA+	4.10					
20	F }	★ ✓	3.35		← 7	.25	F	BA, EIA how much later shared w husband
21	M }	< S, ✓ SA ow	3.25		← 8			
22	F	★ SA	.35					
23	F	★ SA+ discuss	1.15		← 9			
24	M }	< S			10 →	.53	M }	EIA, SA ow
25	F }	< S				.53	F }	joined M at SA+BA only discuss BA
26	M }	>						
27	F }	>				.20	F	SA
28	F }	>				✓	M }	
29	F }	>				✓	F }	
30	M	★ SA only	1.35			✓	M }	
31	M }	★ look, SB	.15			✓	F }	
32	Cf }	SB						
33	M }	★ SA	.30			✓	M	EIA
34	F }	< S						
35	F	↑				✓	M	
36	M	SB						
37	M	↑						
38	M	★ SA+	6.30	look closely at objects	← 11			
39	F	★ BA, SA+	1.37		← 12			
40	M }	★ SA ow, exp LIA	.40					
41	C }	★						
42	C }	★						
43	M }	>						
44	F }	>						
45	C }	>						

Terms and abbreviations in the Observation Sheet which require clarification:

DAILY TEST NUMBERS- indicates the identification numbers of the questionnaires which were given to visitors on that day.

DAILY VISITOR NUMBER- is the ordinal number assigned to each visitor who entered the gallery. The final visitor number showed the total number of visitors who had been observed on that day. Additional sheets were supplied which listed numbers above 45.

SEX- M= male, F= female, C= child, with a further designation of M or F.

ENTRY GROUP- the brackets indicate the size and sexual, as well as adult-child, composition of groups as they entered the first floor gallery.

INITIAL DIRECTION- key to the hieroglyphics which show the direction of visitors' initial movements after entering the gallery and the display cases which attracted them. Sections (Ages) within the experimental display case which atracted visitors, and the sequence in which they viewed these "Ages" are also noted either under this column or under COMMENTS.

→ =to the right.
← =to the left.
↙ =to the rear left.
↑ =down the middle of the gallery.

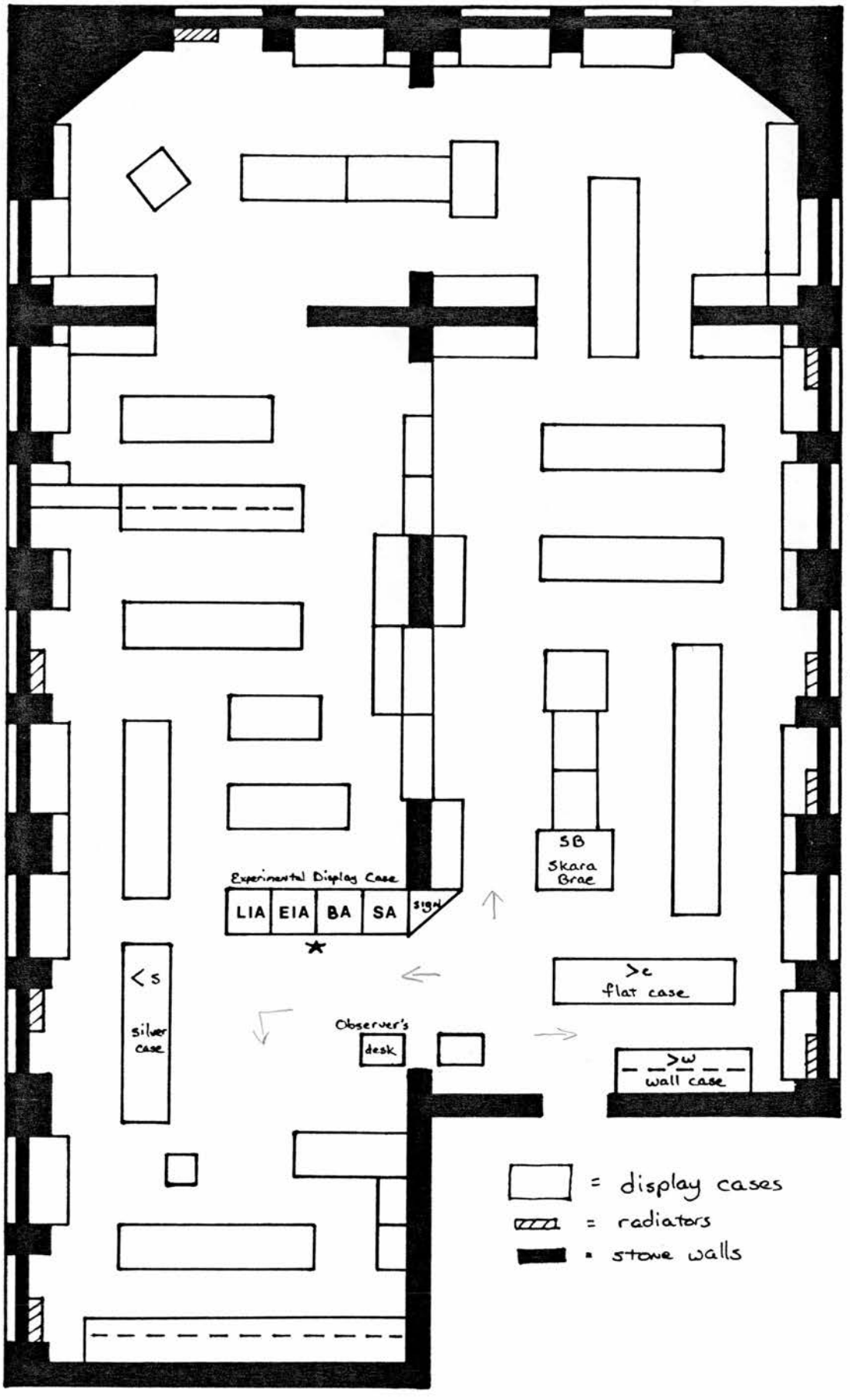
★ =straight to the experimental display case.
>w =to the wall case to the right of the door.
>c =to the flat (horizontal) display case, ahead, and to the right of the door.
SB =to the flat Skara Brae case, straight ahead of the door.
<s =to the silver exhibit in the upright case to the left of the door.

SA =Stone Age. SA+ (or on) = SA thru LIA in correct sequence.
BA =Bronze Age.
EIA =Early Iron Age.
LIA =Late Iron Age.
✓ =glanced at the sign, 2 seconds or less.

VIEW TIME- length of time, in minutes and seconds, spent looking at the experimental display case.
✓ = glance at the experimental display case for 2 seconds or less.

TEST # - identity number of the questionnaire given to the adjacent visitor.
The arrow indicates the visitor tested.

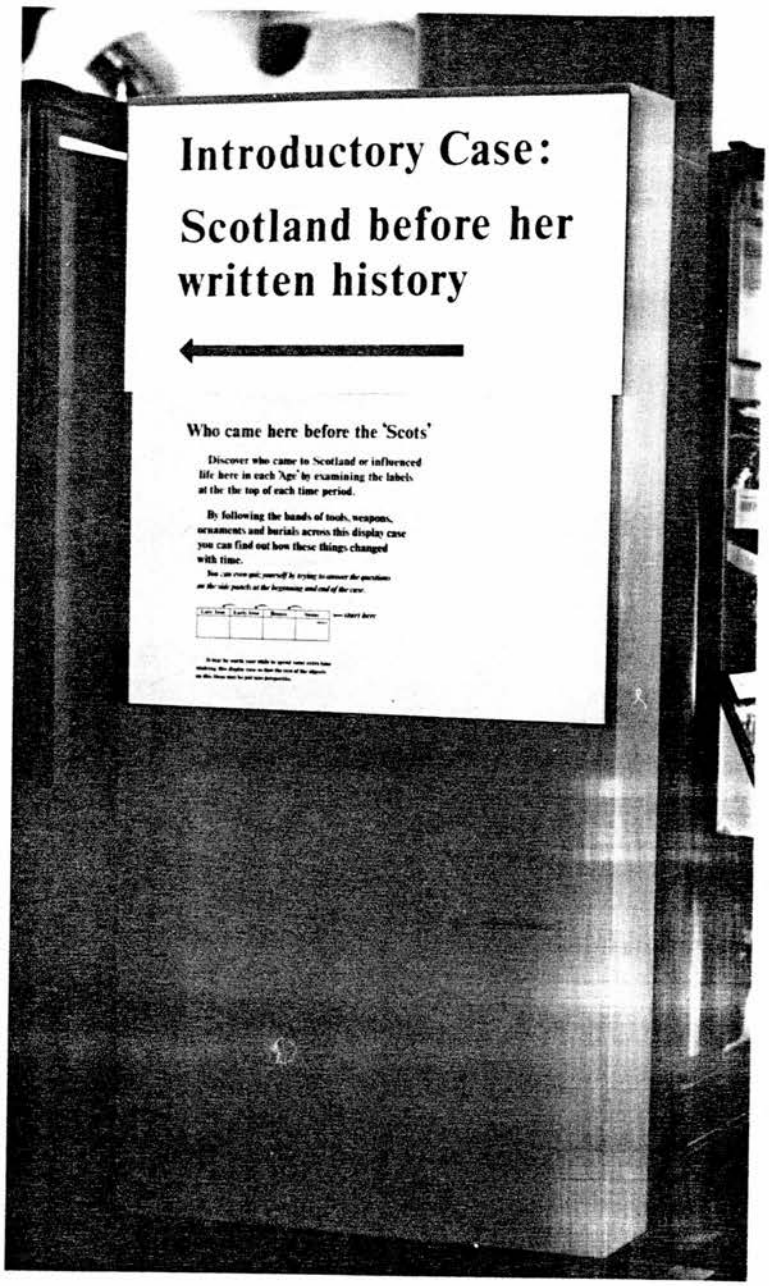
PLAN OF THE MUSEUM FLOOR



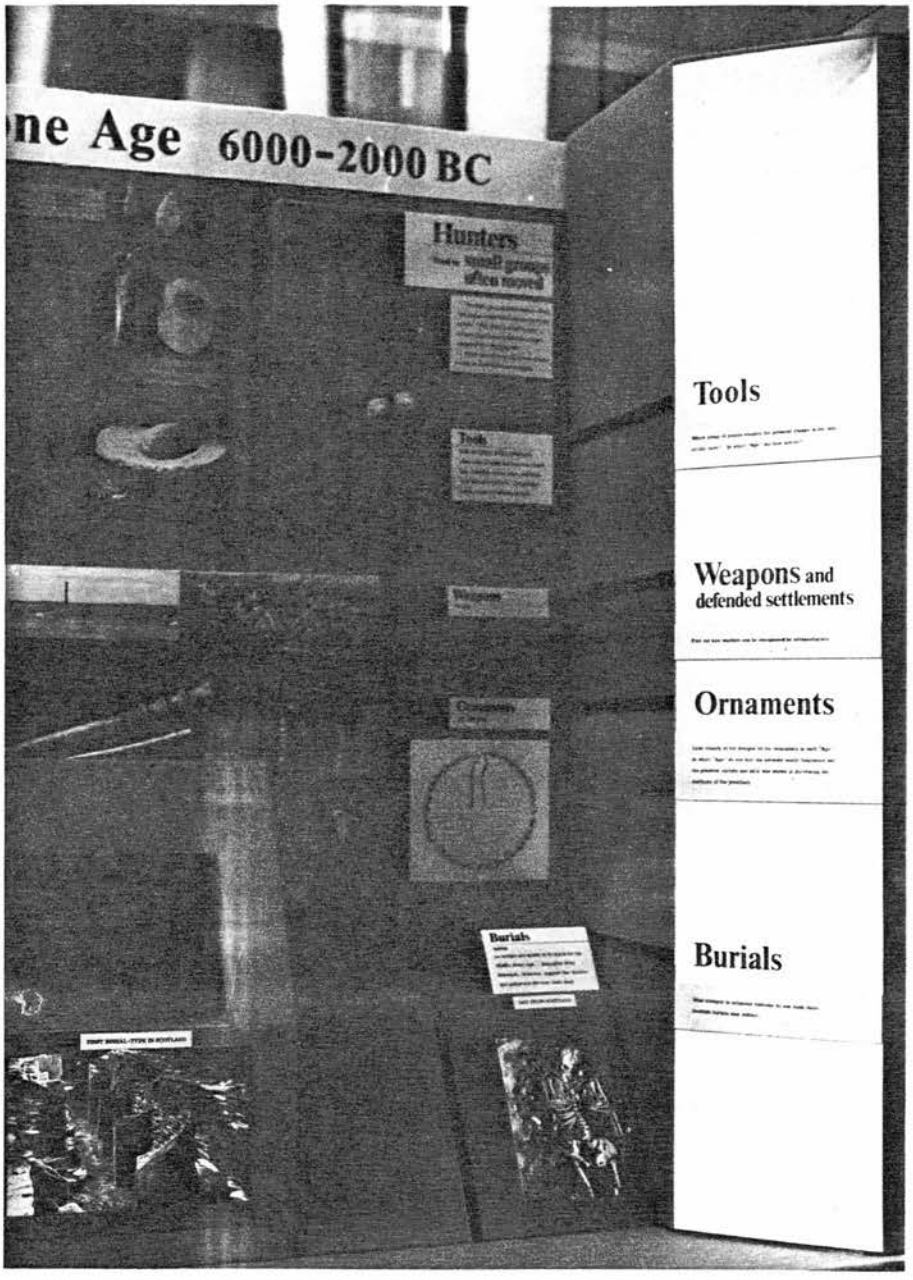
Appendix B. Photographs of Experimental Display Treatments:

Appendix B.I. Goal Photographs:

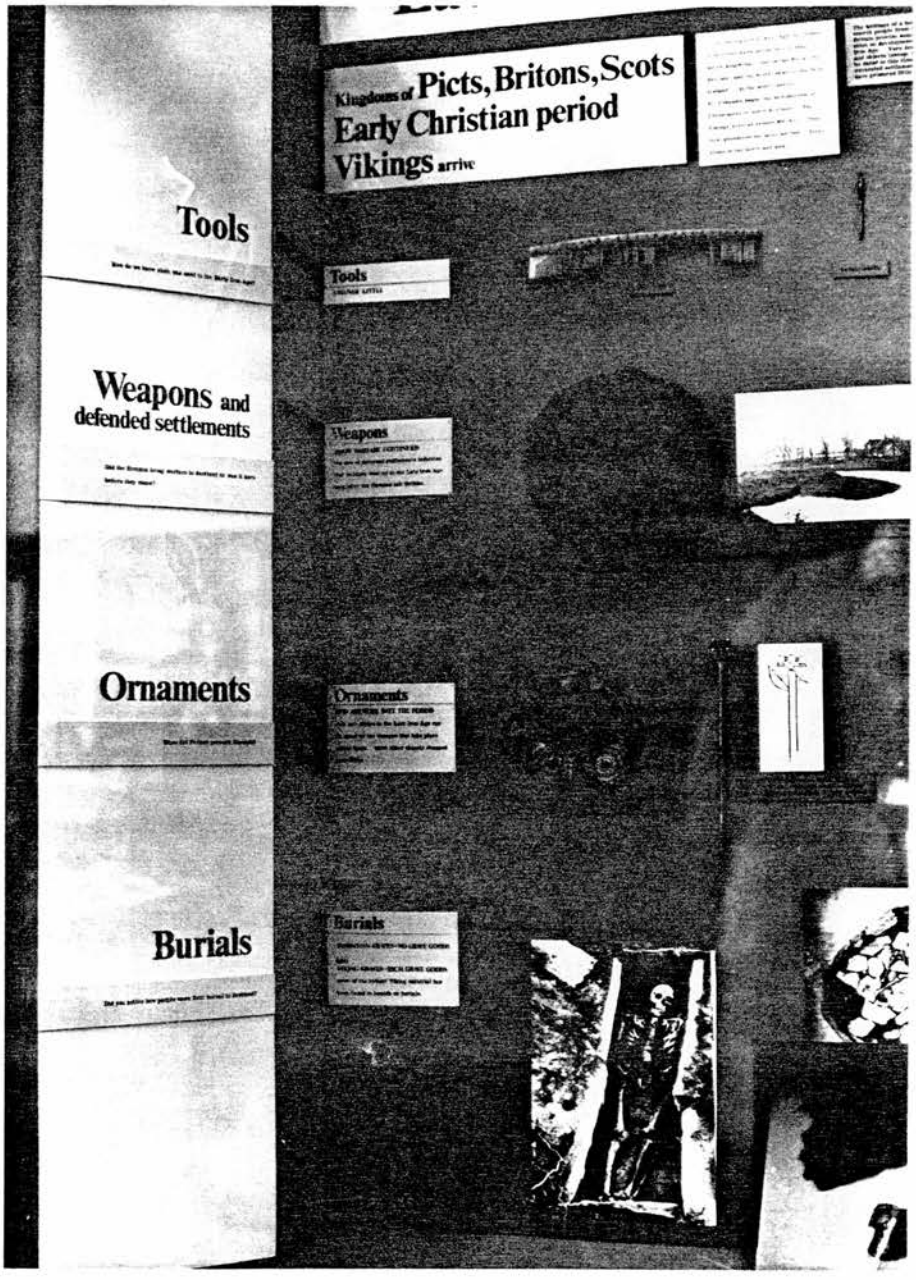
SIGN BESIDE THE DISPLAY CASE.
NOTE THE GOAL STATEMENTS BELOW THE ARROW.



GOAL STATEMENTS (FAINT) AT THE BEGINNING OF THE DISPLAY CASE, I.E. THE RIGHT-HAND SIDE. TO READ THEM SEE APPENDIX C.III.

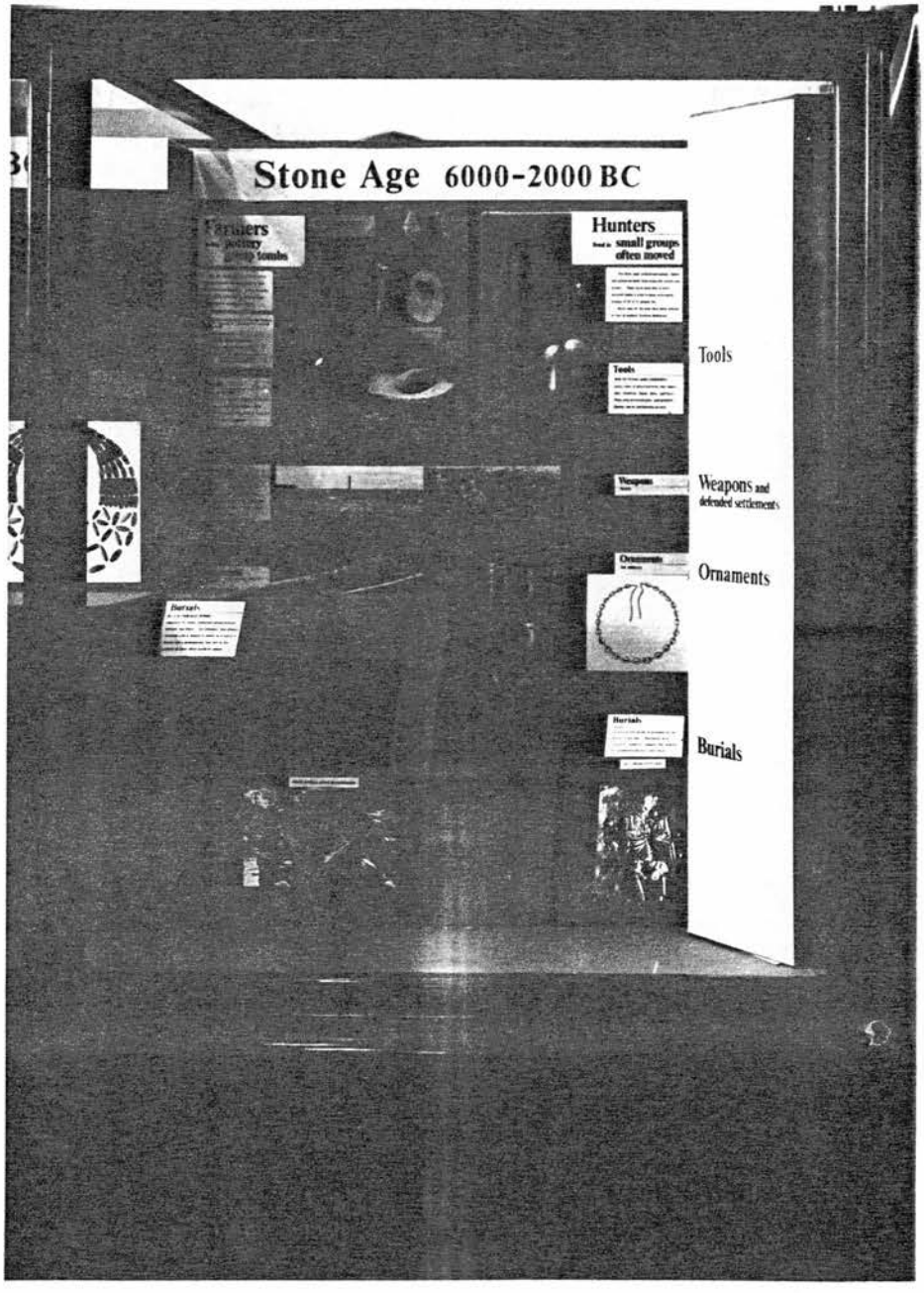


GOAL STATEMENTS AT THE END OF THE DISPLAY CASE.

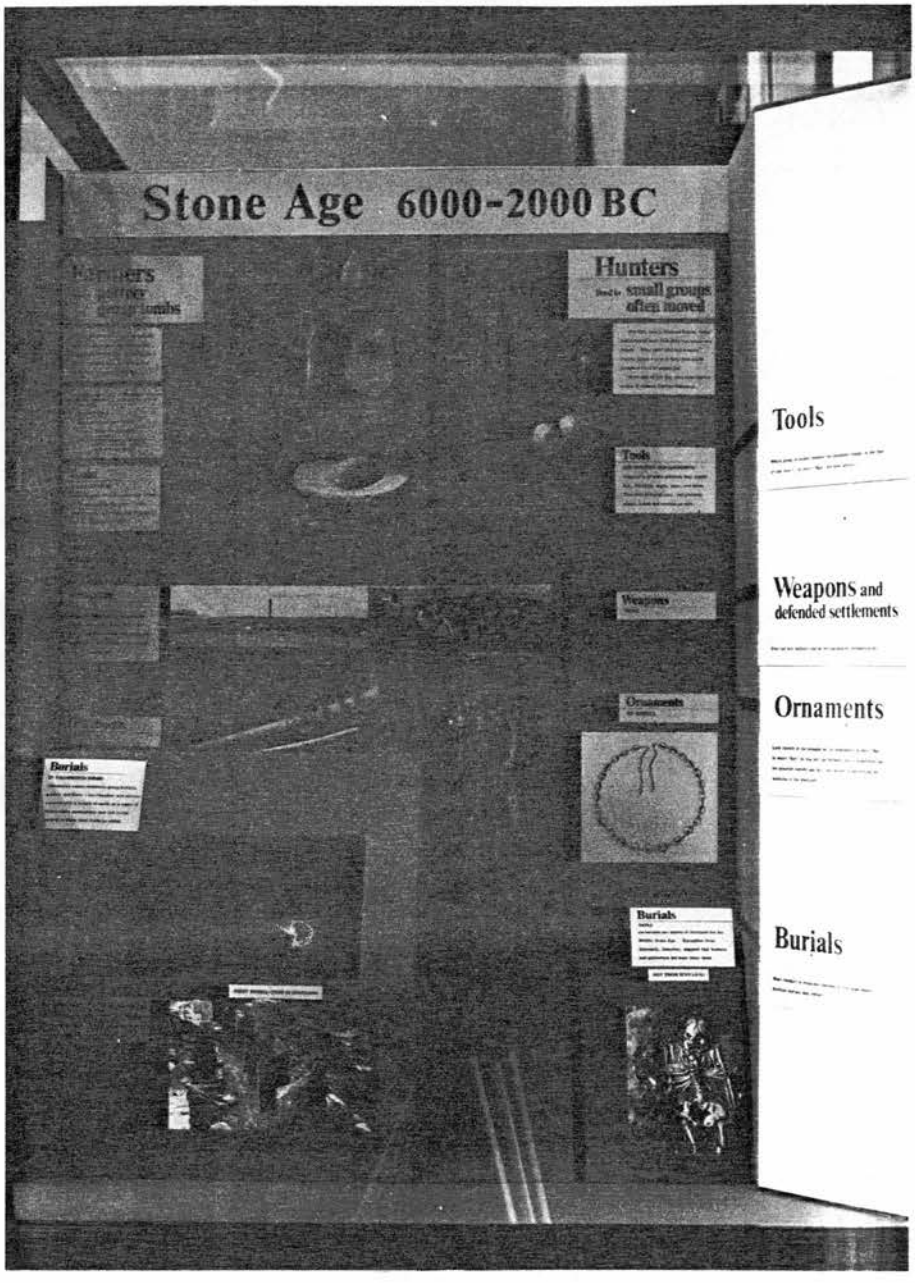


Appendix B.II. MT and MTG Treatment Photographs:

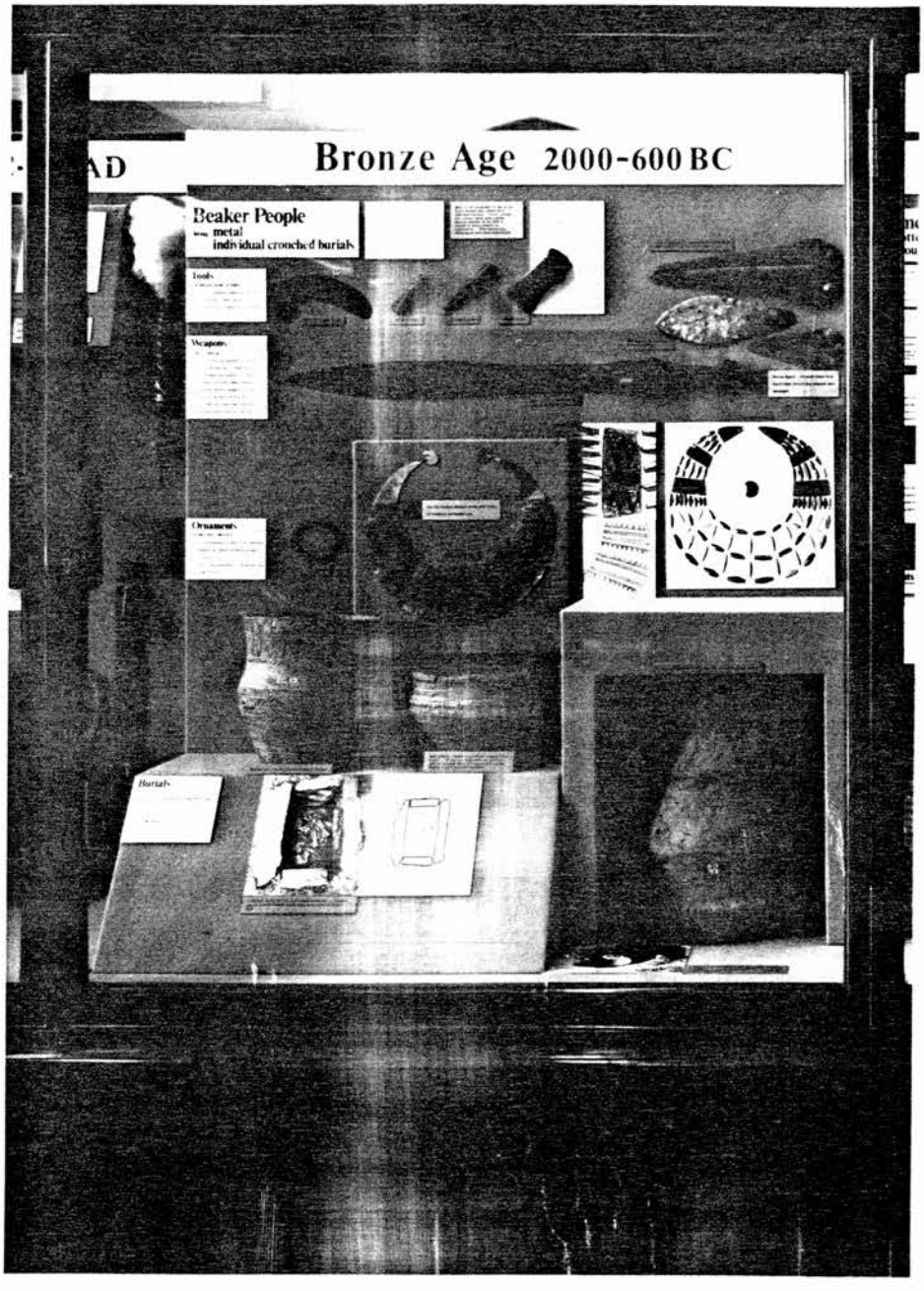
STONE AGE WITHOUT GOALS
MT



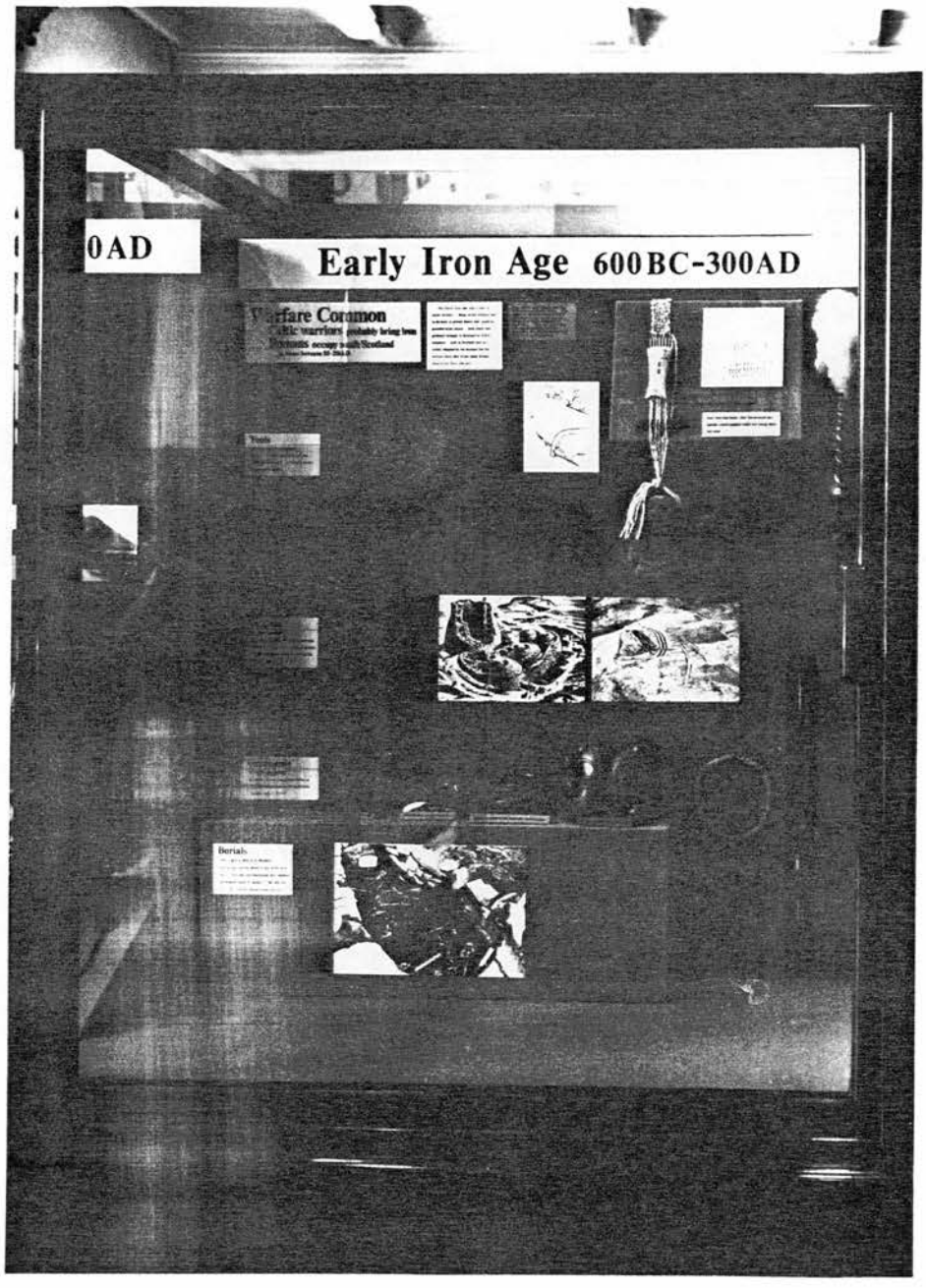
STONE AGE
MTG
MULTI-TRACK OBJECTS AND LABELS, WITH GOALS



BRONZE AGE
MT OR MTG
(SINCE GOALS ONLY APPEAR AT THE BEGINNING AND END OF THE DISPLAY CASE, THE AGES IN THE MIDDLE, I.E. BRONZE AGE AND EARLY IRON AGE, APPEAR THE SAME WHETHER OR NOT GOALS ARE PRESENT OR ABSENT.)

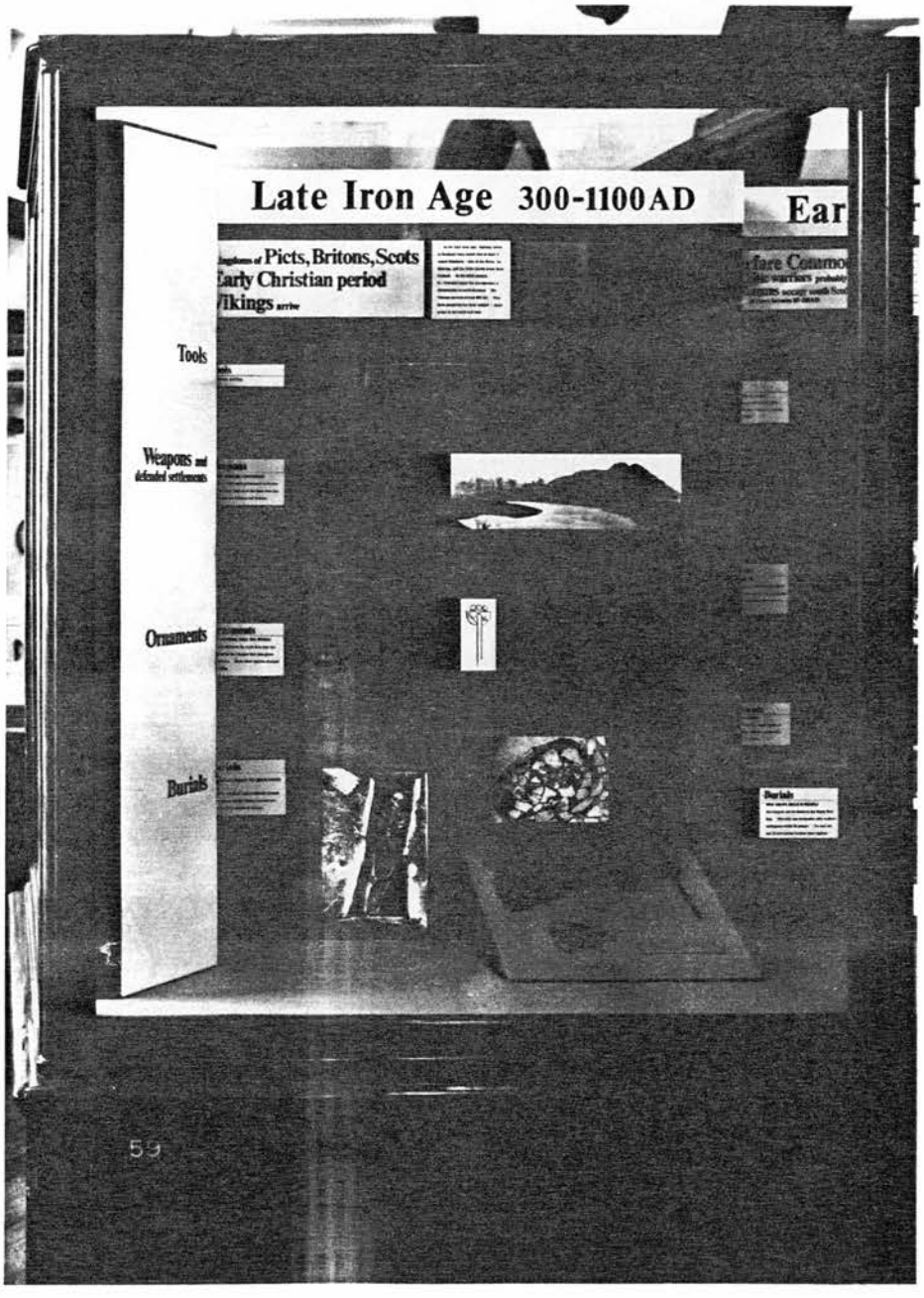


EARLY IRON AGE
MT OR MTG

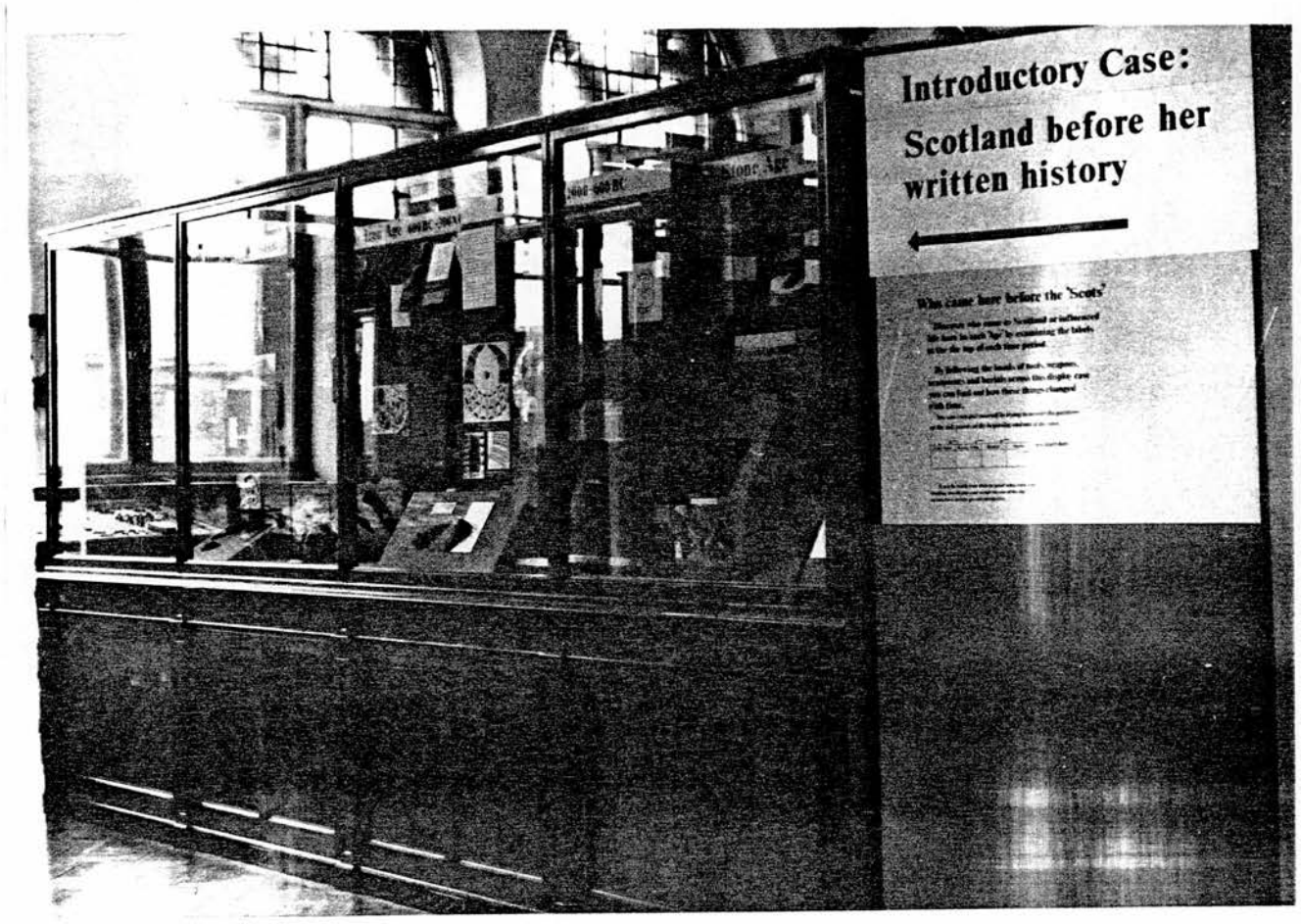


EARLY IRON AGE
MT OR MTG
SHOWING OBJECTS IN USE, I.E.
WEAVING COMB AND LOOM WEIGHT, TOP MIDDLE;
SPINDLE WHORL SPINNING THREAD, TOP RIGHT;
ORNAMENTAL PINS ATTACHED TO VERTICAL RIBBON, BOTTOM RIGHT.

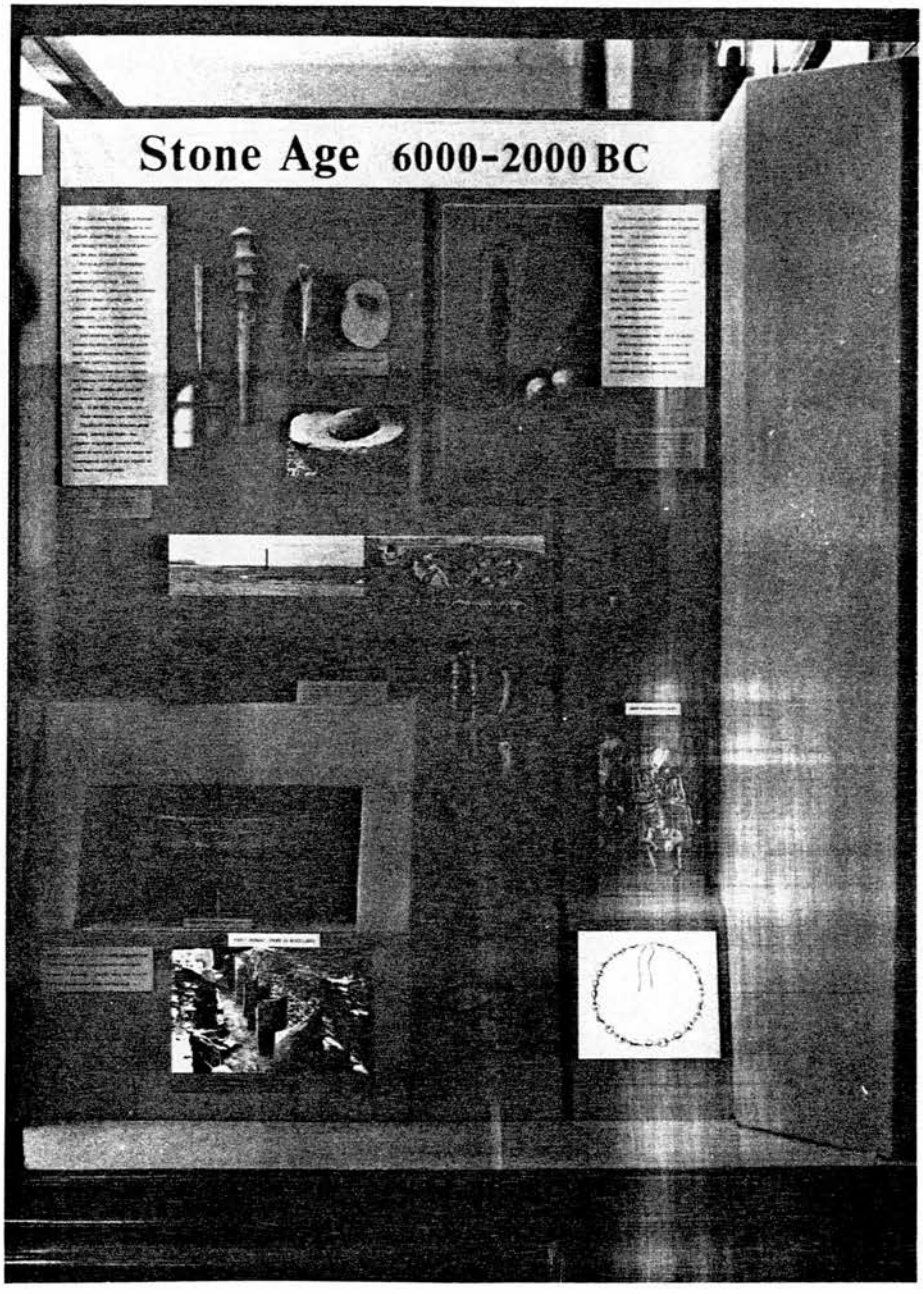
LATE IRON AGE
MT



Appendix B.III. BU and BUG Treatment Photographs:



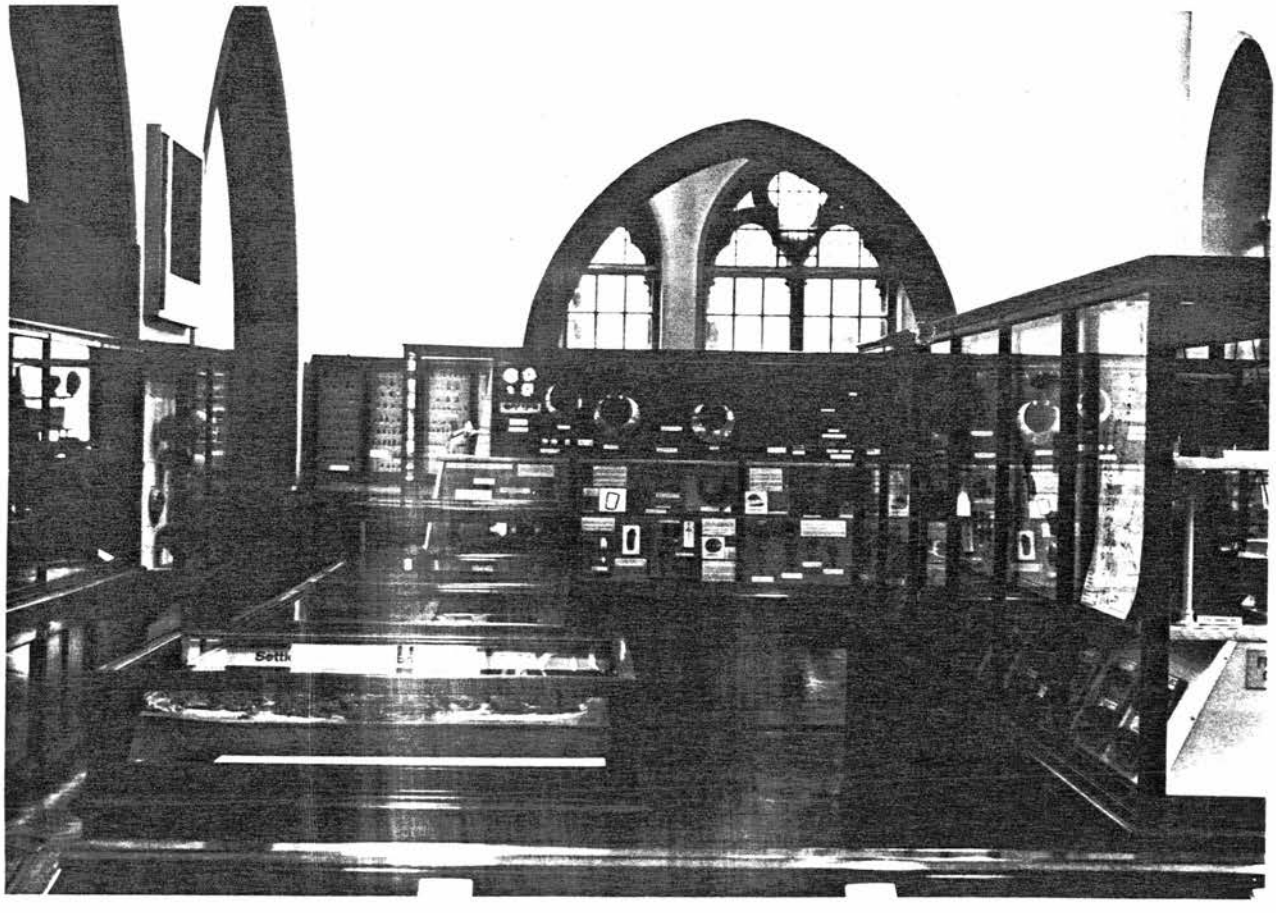
STONE AGE
BU

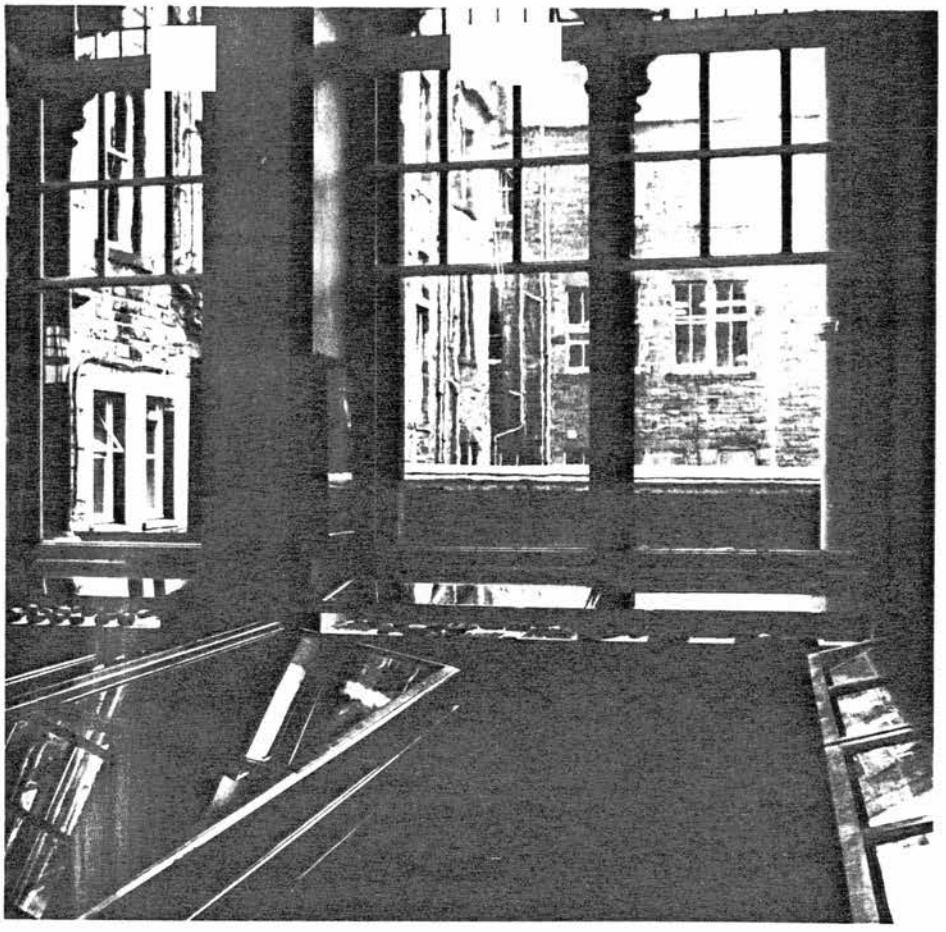


Appendix B.IV. Views from the Gallery Entrance:

1. VIEW TO THE LEFT OF THE ENTRANCE TO THE FIRST FLOOR.
(MTA TREATMENT IN THE DISPLAY CASE)

2. VIEW STRAIGHT AHEAD OF THE ENTRANCE TO THE FIRST FLOOR GALLERY

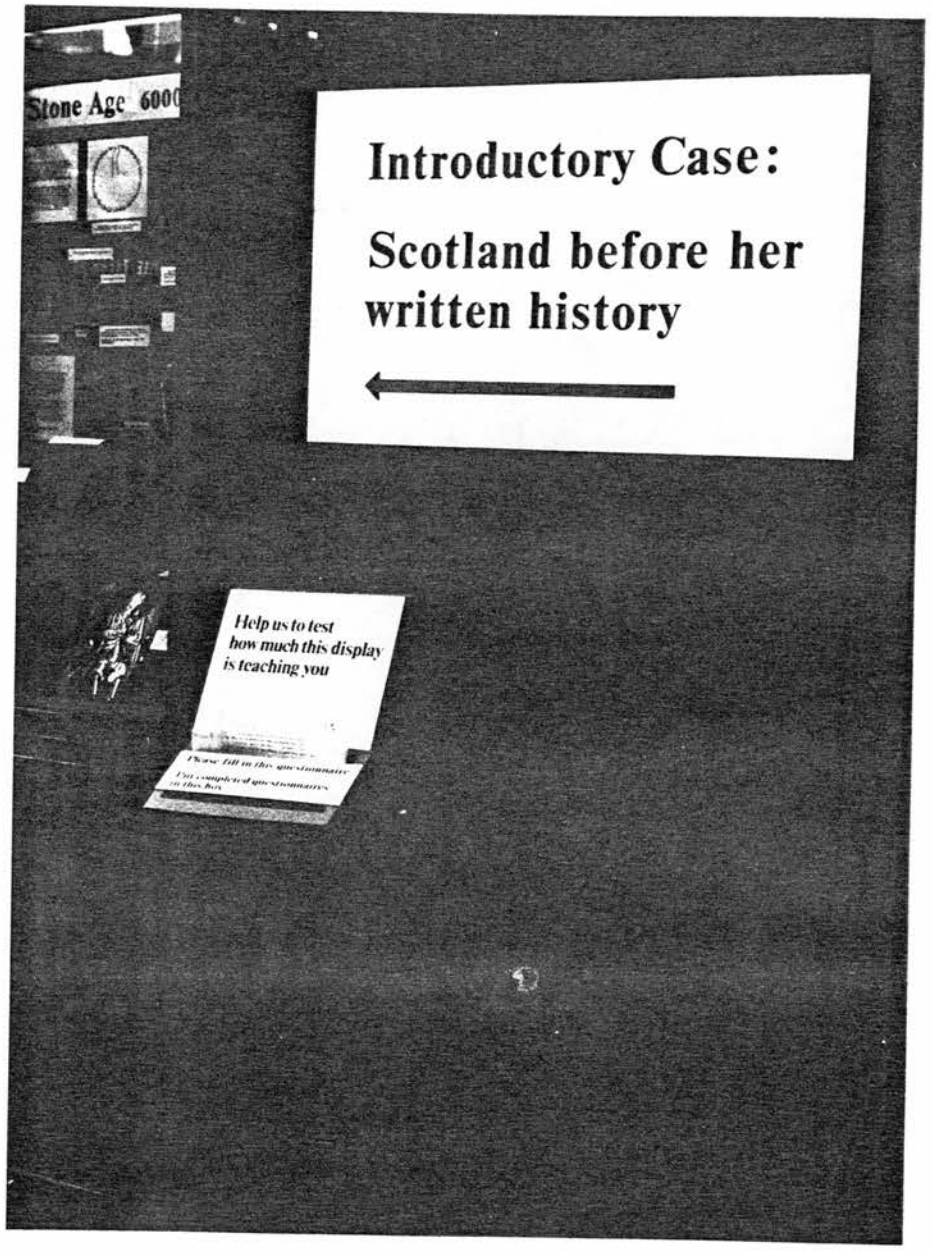




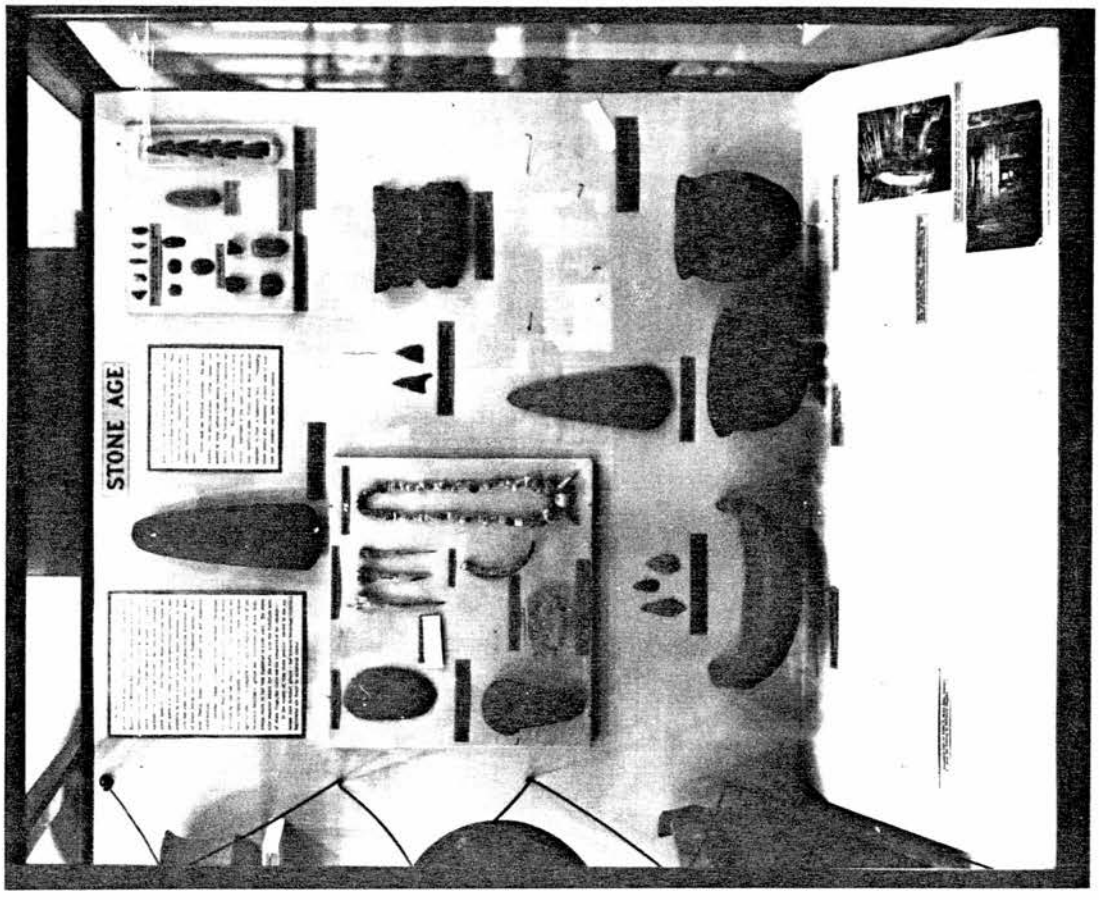
1. VIEW TO THE RIGHT OF THE ENTRANCE TO THE FIRST FLOOR

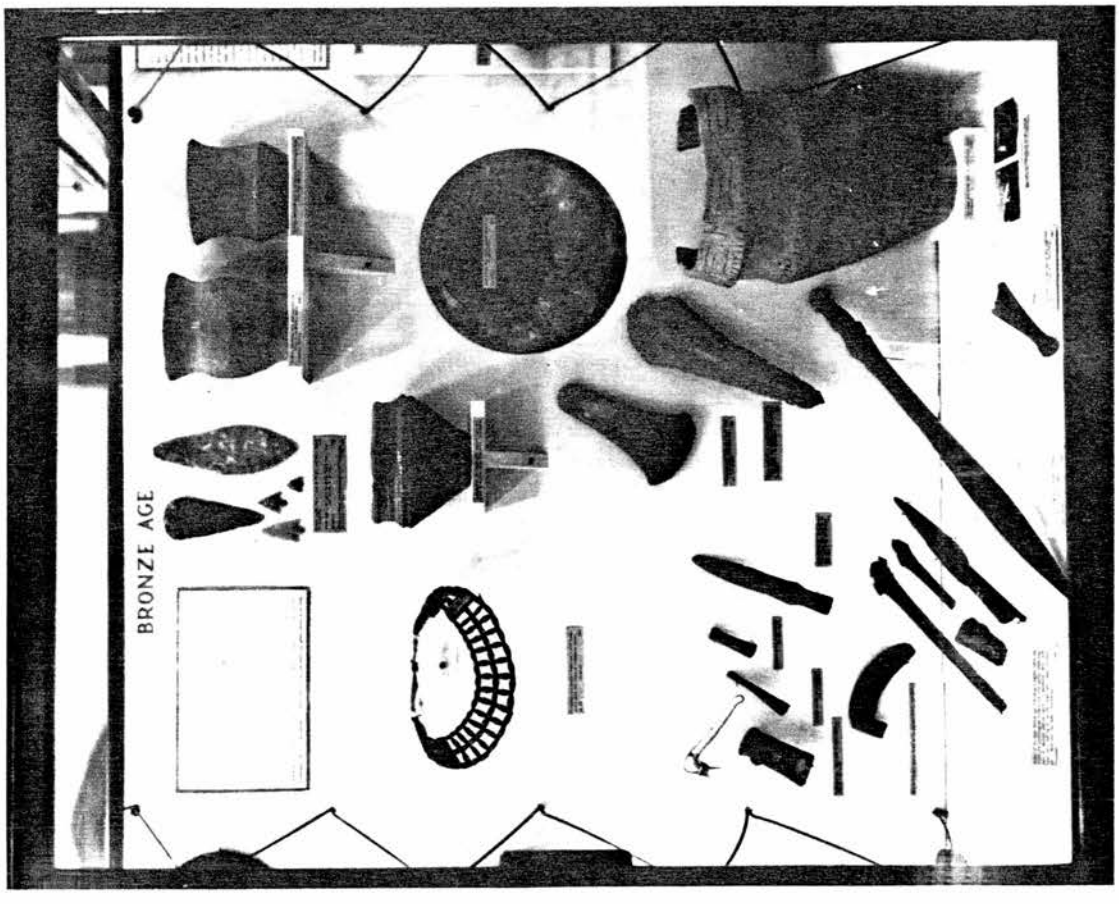
2. VIEW OF THE OBSERVOR, LEFT OF THE ENTRANCE (PILOT DISPLAY)

Appendix B.V. Test Yourself Box Photograph

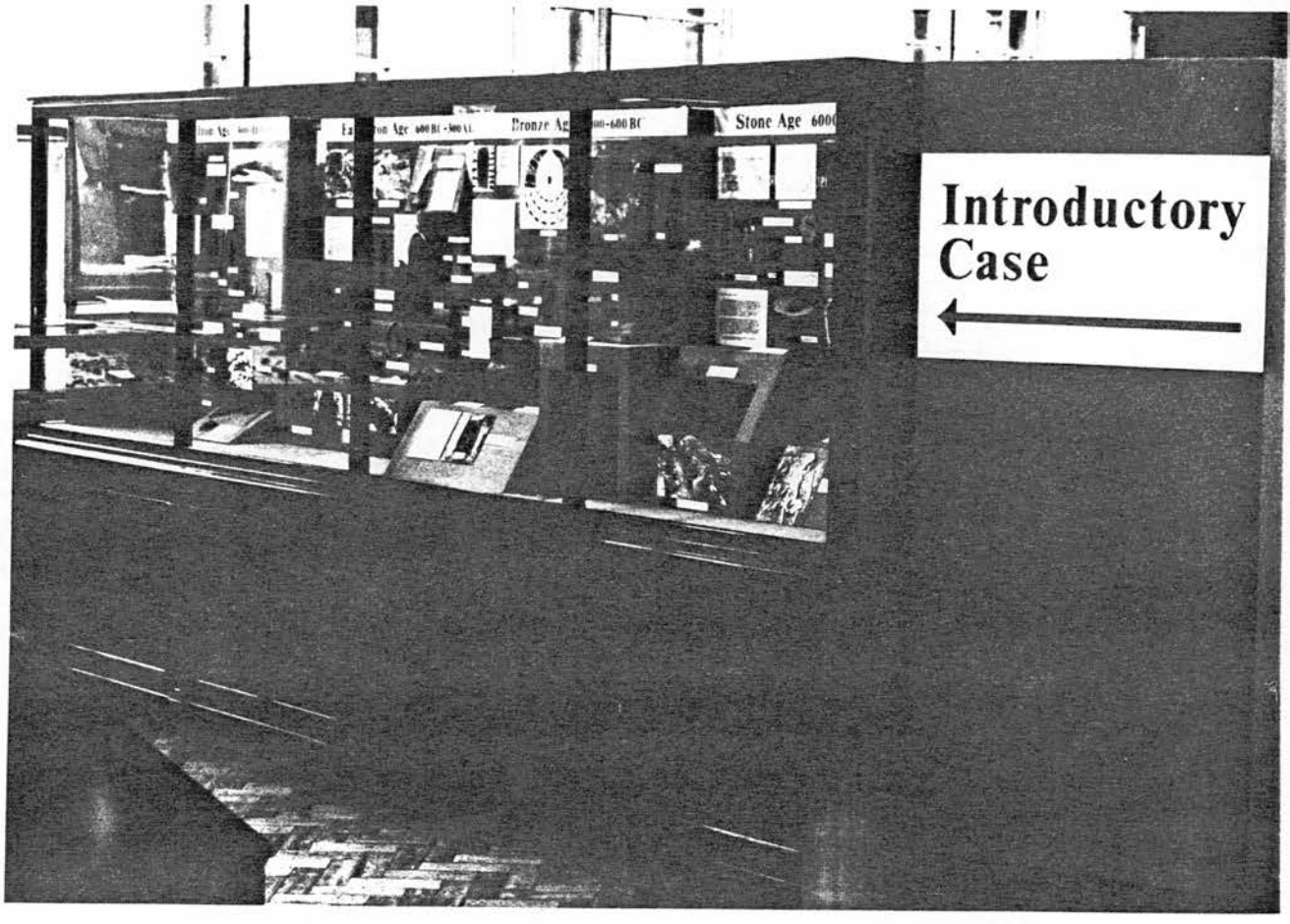


Appendix B.VI. Control Study Photographs– Original Yellow Case

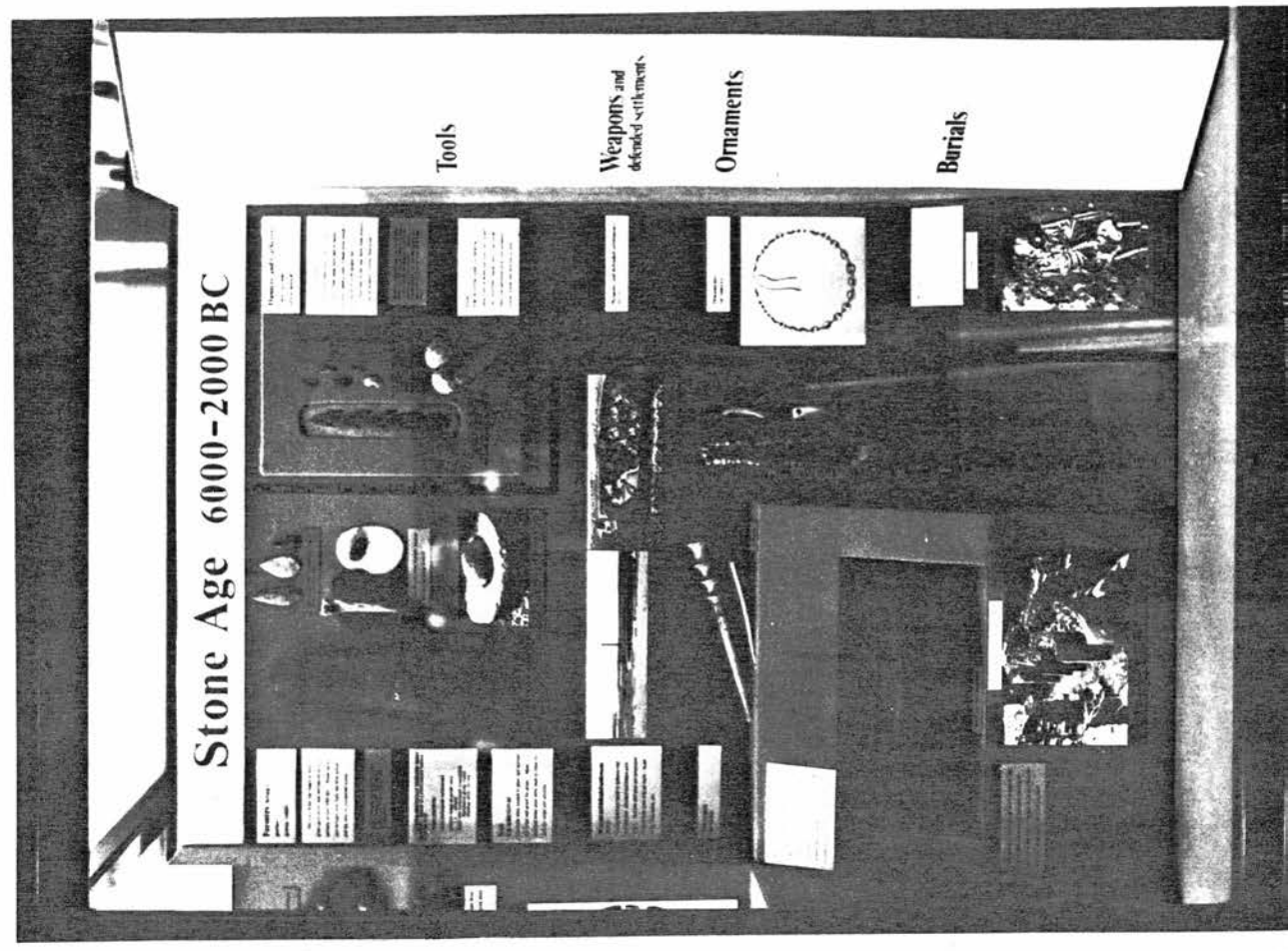


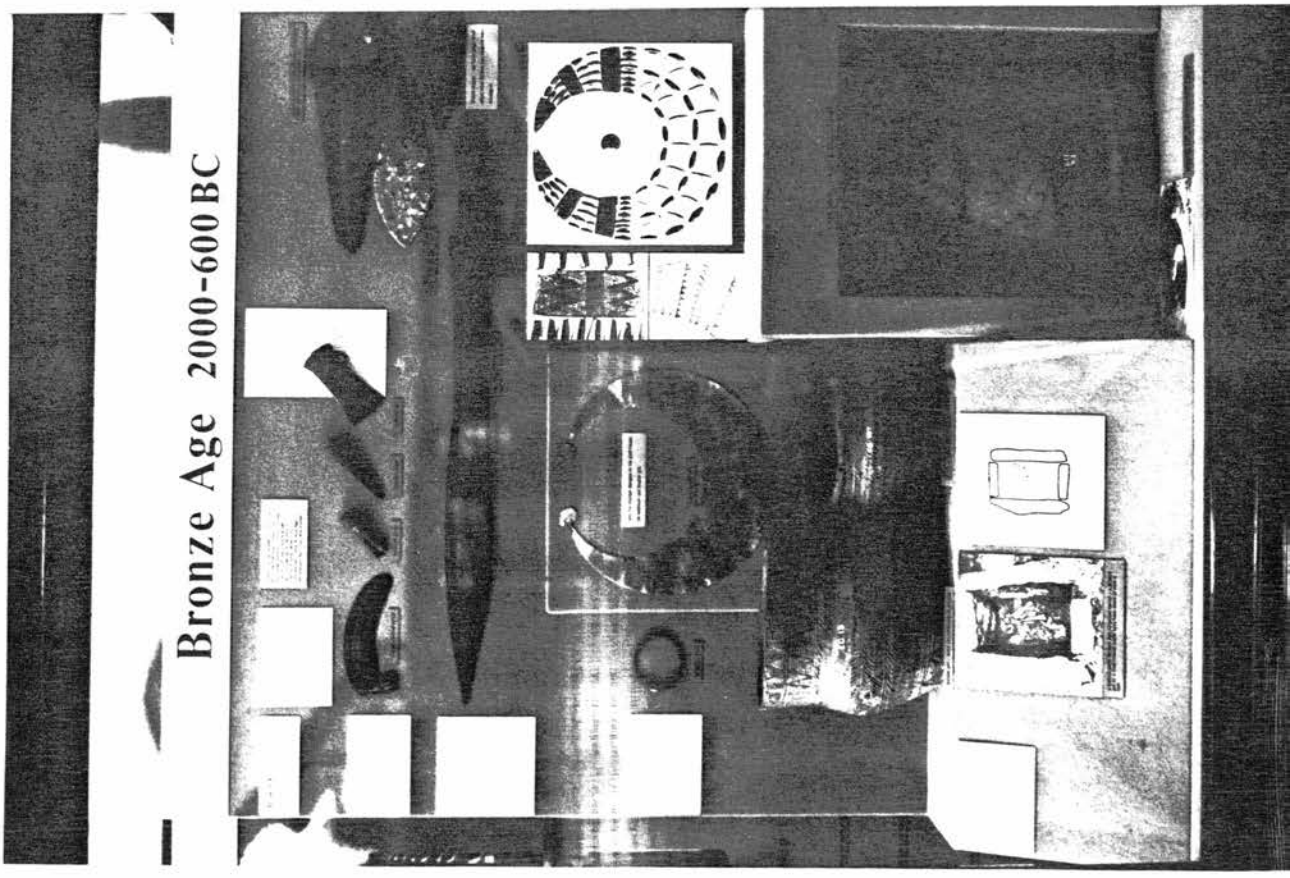


Appendix B.VII. Control Study Photographs– Green Pilot Case



Appendix B.VIII. Control Study Photographs– MTa "small headings"





Appendix B.IX. Control Study Photographs– Alleval Display, containing MT, BT, and BU in one display case.

Appendix C. Labels- Multi-track and Block -used in Main Display Treatments:

Appendix C.I. MT and MTG "Multi-Track" Labels

STONE AGE, MT:

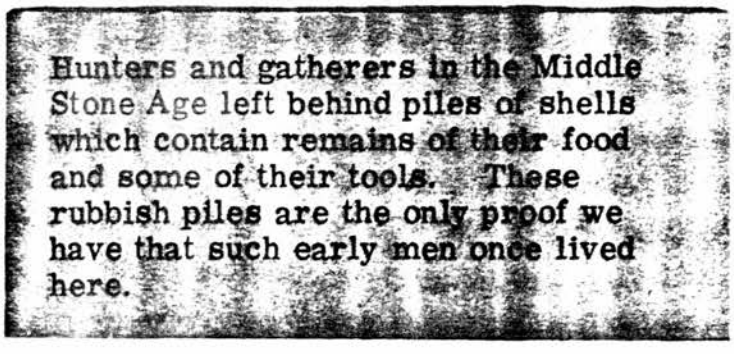
Hunters and Gatherers

small groups

often moved

The first men in Scotland hunted, fished and gathered their food along the coasts and rivers. They must have had to move several times a year to keep even small groups of 15 to 20 people fed.

Their way of life may have been similar to that of modern Caribou Eskimoes.



Hunters and gatherers in the Middle Stone Age left behind piles of shells which contain remains of their food and some of their tools. These rubbish piles are the only proof we have that such early men once lived here.

STONE AGE, MT:

Tools

FOR HUNTING AND GATHERING

using tools of stone and bone they caught fish, shellfish, seals, deer, and birds. They also gathered nuts, and probably plants, seeds and berries as well.

Weapons and defended settlements

NONE

Ornaments

OF BONE

Burials

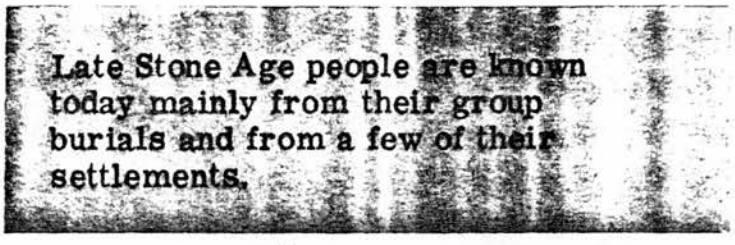
NONE

no burials are known in Scotland for the Middle Stone Age. Examples from Denmark, however, suggest that hunters and gatherers did bury their dead.

STONE AGE, MT:

Farmers bring
pottery
group tombs

The Late Stone Age began in Scotland when agriculture was introduced by new settlers around 3500 BC. These farmers also brought with them the first pottery and the idea of chambered tombs.



Farming gradually brought about these changes - (based on a more secure method of getting food)

larger population

more permanent settlements

greater range of goods used (example - pottery)

large ceremonial monuments built

i. e. - chambered group tombs,

and later, standing stone circles.

STONE AGE, MT:

Tools

FOR AGRICULTURE

new tools were needed to plant and harvest the crops and grind the grain. Many polished stone axes were used to clean the land for crops and animals.

Weapons and defended settlements

NONE

nothing has been found to indicate that humans even attacked and killed each other. Arrows and axes are presumed to have been used only as tools - to get food, chop meat, etc.

Ornaments

OF SHELLS

Burials

IN CHAMBERED TOMBS

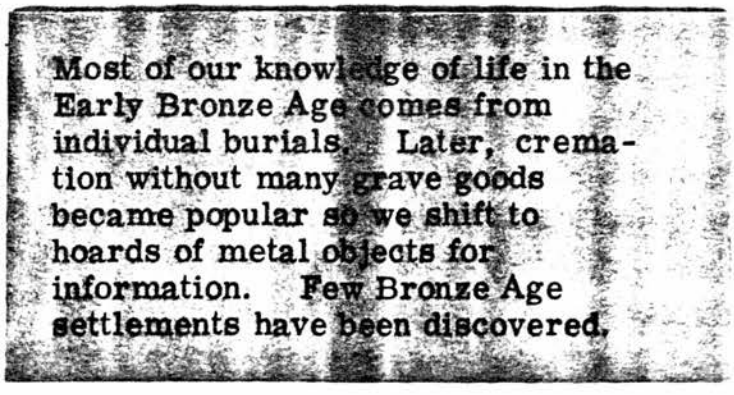
the building of large group tombs suggests that new religious and social patterns developed because of farming. People were now free to work on things apart from getting food.

chambered tombs contained group burials, pottery and flints-the chamber was always covered with a mound of earth or a cairn of stones and a passageway was left to the outside so later dead could be added.

BRONZE AGE, MT:

Beaker People bring metal individual crouched burials

The magic of metal-working was brought to Scotland by Beaker people from the Low Countries and north Germany. Along with copper, bronze and gold, Beaker groups also brought the Beaker pots from which they are named and the use of crouched burials.



BRONZE AGE, MT:

Tools

OF BRONZE RARE AT FIRST

stone tools remained important until the late Bronze Age - when the ability to make bronze out of copper and tin became widespread.

Weapons and defended settlements

FIRST APPEAR

warfare is first recognizable at the very end of the Bronze Age. Then the first known defended settlements (hill forts) are present along with many weapons.

Although definite weapons are first found in the Early Bronze Age - they may have been used more for personal status than for fighting.

BRONZE AGE, MT:

Ornaments

OF PRECIOUS METALS

skilled workmanship in objects for personal use suggest that wealth or status groups were formed.

most surviving ornaments in the Bronze Age are made of gold.

Burials

first
CROUCHED IN STONE COFFINS (cists)
often with a Beaker or a Food Vessel pot.

later CREMATED
ashes of dead put in cinerary urns which were buried upside-down.

EARLY IRON AGE, MT:

Warfare Common

Celtic warriors probably bring iron

Romans occupy south Scotland at times between 85-211AD

The Early Iron Age was a time of much warfare. Many of the villages had to be built in places where they could be guarded from attack. Iron itself was probably brought to Scotland by Celtic invaders. Life in Scotland was not really changed by the Romans but the natives were able to get many Roman objects for their own use.

Iron objects don't usually last long in the archaeological record. Thus they are mainly known in Scotland from a few scrap hoards (two of which were found inside bronze cauldrons). Although many Early Iron Age settlements have been excavated, few notable objects have been found in most of them.

EARLY IRON AGE, MT:

Tools

OF IRON MORE COMMON

iron ore was much easier to find than materials for bronze - so its use spread to all aspects of life.

Weapons and defended settlements

INDICATE WARFARE

many large defended hill-forts and brochs (tower-like farmhouses) as well as smaller forts and farms have been found from the Early Iron Age.

Ornaments

BRONZE PREFERRED

the ability to cast iron wasn't known at this time so finer decorated metalwork was done only in bronze.

Burials

ONE GRAVE HELD 21 PEOPLE

few burials can be dated to the Early Iron Age. The only one excavated with modern techniques holds 21 people. No one can say if such group burials were typical or not.

LATE IRON AGE, MT:

Kingdoms of Picts, Britons, Scots
Early Christian Period
Vikings arrive

In the Late Iron Age, fighting tribes in Scotland were united into at least 3 major kingdoms - that of the Picts, the Britons, and the Scots (newly come from Ireland). In the sixth century, St. Columba began the introduction of Christianity to north Scotland. The Vikings arrived around 800 AD. They first plundered but later settled in large areas in the north and west.

The writings of a few Romans and church people from other parts of Britain provide some of our information on developments in the Late Iron Age. Very few settlements and objects (except for art-work) can be dated to this time. Even the excavated settlements we know of have produced little unusual material.

LATE IRON AGE, MT:

Tools

CHANGE LITTLE

Weapons and defended settlements

SHOW WARFARE CONTINUED

the use of defended settlements indicates that warfare went on in the Late Iron Age long after the Romans left Britain.

Ornaments

AND ARTWORK DATE THE PERIOD

only art-styles in the Late Iron Age can be dated by the changes that take place within them. Most other objects changed very little.

Burials

CHRISTIAN GRAVES-NO GRAVE GOODS

later
VIKING GRAVES-RICH GRAVE GOODS

most of the richest Viking material has been found in hoards or burials.

EXAMPLES OF STONE AND BRONZE AGE TERTIARY IDENTITY LABELS:

bone beads and perforated tusks

bone beads and a tusk

found in a chambered tomb

NOT FROM SCOTLAND

bone weaving comb

clay loomweight

spindle whorl

bronze and bone pins

FIRST BURIAL-TYPE IN SCOTLAND

bone awl and chert scraper suggest leather was being worked.

bone weaving comb, clay loomweight and spindle whorl suggest cloth was being made and used.

bronze dagger - the basic shape from which other Bronze Age weapons were developed.

note the similar designs on the gold lunula, jet necklace and Beaker pot.

Appendix C.II. BU and BUG, BT and BTG, "Block" Labels

STONE AGE (REDUCED), BU-BT:

The Late Stone Age began in Scotland when agriculture was introduced by new settlers around 3500 BC. These farmers also brought with them the first pottery and the idea of chambered tombs.

Farming gradually created these changes - (based on a more secure method of getting food): a larger population; more permanent settlements; a greater range of goods used, i.e. - pottery; and large new ceremonial monuments; i.e. - chambered group tombs, and standing stone circles.

New tools were needed to plant and harvest the crops and grind the grain. Many polished stone axes were used to clear the land for crops and animals.

Nothing has been found to indicate that humans ever attacked and killed each other. Arrows and axes are presumed to have been used only as tools - to get food, chop meat, etc.

Their ornaments were made of bone.

Chambered tombs contained group burials, pottery and flints - the chamber was always covered with a mound of earth or a cairn of stones and a passageway was left to the outside so later dead could be added.

The first men in Scotland hunted, fished and gathered their food along the coasts and rivers. They must have had to move several times a year to keep even small groups of 15 to 20 people fed. Their way of life may have been similar to that of modern Caribou Eskimoes.

Using tools of stone and bone they caught fish, shellfish, seals, deer, and birds. They also gathered nuts, and probably plants, seeds and berries as well.

No evidence of weapons or of defended settlements has been found.

Their ornaments were made of shells.

No burials are known in Scotland for the Middle Stone Age. Examples from Denmark, however, suggest that hunters and gatherers did bury their dead.

BRONZE AGE, BU-BT:

The magic of metal-working was brought to Scotland by Beaker people from the Low Countries and north Germany. Along with copper, bronze and gold, Beaker groups also brought the Beaker pots from which they are named and the use of crouched burials.

Stone tools remained important until the late Bronze Age - when the ability to make bronze out of copper and tin became widespread.

Warfare is first recognizable at the very end of the Bronze Age. Then the first known defended settlements (hill forts) are present along with many weapons. Although definite weapons are first found in the Early Bronze Age - they may have been used more for personal status than for fighting.

Skilled workmanship in objects for personal use suggests that wealth or status groups were formed. Most surviving ornaments in the Bronze Age are made of gold.

Burials were first crouched in stone coffins (cists) often with a Beaker or a Food Vessel pot. Later, ashes of the dead were put in cinerary urns which were buried upside-down.

EARLY IRON AGE, BU-BT:

The Early Iron Age was a time of much warfare. Many of the villages had to be built in places where they could be guarded from attack. Iron itself was probably brought to Scotland by Celtic invaders. Life in Scotland was not really changed by the Romans but the natives were able to get many Roman objects for their own use.

Tools of iron were more common. Iron ore was much easier to find than materials for bronze - so its use spread to all aspects of life.

Weapons and defended settlements indicate warfare. Many large defended hill-forts and brochs (tower-like farmhouses) as well as smaller forts and farms have been found from the Early Iron Age.

The ability to cast iron wasn't known at this time so finer decorative metalwork was done only in bronze.

Few burials can be dated to the Early Iron Age. The only one excavated with modern techniques holds 21 people. No one can say if such group burials were typical or not.

LATE IRON AGE, BU-BT:

In the Late Iron Age, fighting tribes in Scotland were united into at least 3 major kingdoms - that of the Picts, the Britons, and the Scots (newly come from Ireland). In the sixth century, St. Columba began the introduction of Christianity to north Scotland. The Vikings arrived around 800 AD. They first plundered but later settled in large areas in the north and west.

Tools show little change.

The use of defended settlements indicates that warfare went on in the Late Iron Age long after the Romans left Britain.

Only art-styles in the Late Iron Age can be dated by the changes that take place within them. Most other objects changed very little.

Christian graves held no grave goods. Later Viking graves, however, contained rich grave goods. Most of the richest Viking material has been found in hoards or burials.

THESE IDENTITY LABELS WERE THE SAME FOR THE MTS, BTS, AND BUS.:

cowrie shells, found with definite holes, may have been strung into a necklace like this by Middle Stone Age people

Skara Brae, Orkney - like other settlements in the Late Stone Age, was undefended. Nothing suggests that warfare was present.

this saddle quern was one of the new farm tools needed to grind grain

hill on which the Dunadd fortress was built

hill fort at Woden Law - an aerial view of the defenses

a plough sock helped protect a wooden plough from wear - fitted to the underside of the 'share'

plan of the hill fort at Dunadd

saddle querns of this type were used to grind grain for over 3,000 years

Appendix C.III. Goal Labels at Beginning and End of Case

Which group of people created the greatest change in the way-of-life here? In which "Age" did they arrive?

Did the Romans bring warfare to Scotland or was it here before they came?

Look closely at the designs on the ornaments in each "Age". In which "Age" do you feel the artwork really flourished and the greatest variety and skill was shown in decorating the surfaces of the jewellery.

What changes in religious customs do you think these Scottish burials may reflect?

How do we know cloth was used in the Early Iron Age?

Find out how warfare can be recognised by archaeologists.

When did Pictish artwork flourish?

Did you notice how people were first buried in Scotland?

Appendix D. Pilot and Control Group Questionaires:

Appendix D.I. Example of One of the Questionnaire Prototypes in the Pilot Study (reduced):

Sat Sun
J F M A ____ 1978

Time _____
Variation A B C D E F

The museum staff are trying to make their displays here more interesting to visitors. Please help us! Your criticisms of this exhibit case will help us choose between different methods of presenting objects and information.

Please tick the answer you choose (✓).
If you do not wish to answer a question just leave it blank.

1. What was the first exhibit case that you studied in this room? Was it-
   a.( ) to the left of the door (as you enter)
   b.( ) to the right of the door
   c.( ) the introductory case
   d.( ) the Skara Brae case
   e.( ) other _____

2. What attracted you to the first case that you studied? Was it-
   a.( ) its location near the door
   b.( ) particular objects in the case
   c.( ) the topic covered
   d.( ) the case design
   e.( ) signs directing you to the case
   f.( ) seeing other people looking at it
   g.( ) your wish to avoid crowding by others
   h.( ) other _____

3. Approximately how long did you look at the contents and labels in this particular introductory case?
   Please estimate the time- _____

4. How would you describe your general knowledge of early Scottish history and/or archaeology? Would you say it is?
   a.( ) good
   b.( ) fair
   c.( ) poor
   d.( ) non-existent

5. What did you see in this particular exhibit case which you feel someone today might find useful or attractive?

6. How many of the labels in this exhibit case did you read?

a.( ) all or almost all of them
b.( ) most of them
c.( ) some of them
d.( ) none or very few of them

7. How much did you learn from this case?

Most of the following questions are rather difficult. If you don't remember the answer from what you saw in this exhibit case please leave that answer blank.

If you remember the answer from this exhibit case please fill the blank in the following sentences using the term-

Stone Age, Bronze Age, Iron Age, or Post Roman

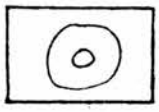
7. The attention paid to decorative war gear suggests that the warrior-class became dominant in Scotland during the _____.

8. Skara Brae is a well known _____site.

9. The _____ period is known primarily from its burials of individuals in which the bodies are doubled up in a stone cist.

10. Pictish carving flourished in the _____.

11. Choose the one correct answer, please



Having seen objects of this shape; you recall that This was used

a.( ) on the end of a harpoon butt
b.( ) to weight the end of a spindle
c.( ) as a mount for a bronze fibula

12. Please tick only the objects below which you recall from the case were representative of the Bronze Age-

a.( ) barbed arrowhead
b.( ) gold lunula
c.( ) beaker
d.( )double-ringed silver chain
e.( )cinerary urn
f.( )penannular brooch

13. Please answer some questions about yourself for our general information:

13. Where are you from?_____

14. Are you ( ) Male or ( ) Female?

15. How often have you visited this museum in the last 12 months?

Please estimate the number of times _____.

16. What is your occupation and what exactly do you do?

_____

17. What is your age group?

( ) 14 or under
( ) 15 to 24
( )25 to 34
( )35 to 49
( )50 to 65
( )66 and over

Thankyou for your help.

Appendix D.II. Control Group– "Test Yourself" Questionnaire (reduced)

S S W F
M A M 9 78

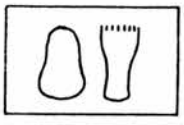
All the answers are in this case
see if you can find them.

no. 513 ty
var. Gt

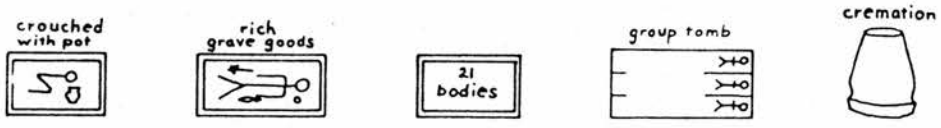
I need your help to find out if this type of display is useful in getting across information. Even if you are not sure of the answer, please try a guess, perhaps you picked up more than you think!

1. Please tick (✓) the one most correct time period for each item on the left.

	Stone Age	Bronze Age	Early Iron Age	Late Iron Age
(example) People first came to Scotland in the-	✓			
Early Christian Period was also the-				
First farmers arrived in the-				
Pictish artwork flourished in the-				
Cremation became popular in the -				
Warfare and Romans were here in the-				

2. 

These objects were both used for one activity-
What was that activity? _____



3. Above are illustrations of early burial types found in Scotland. Please try to put them in order from 1 to 5 - starting with 1 below the first type, and ending with 5 below the last type shown.

4. Please (circle) the two defended sites which were shown in this display case.

Skara Brae Jarlshof Dunadd Mousa

5. What did you find in this case which particularly interested you-

Please (circle) the two objects which had similar designs on them-


gold lunula


bronze dagger


penannular brooch


Kirkwall Type-A pot


jet spacer-plate necklace


halberd


bronze armlet

7. Last one! Try your skill at matching-

Please use a line to connect each item on the left with the main way you think each one changed between the Late Stone Age and the Late Iron Age-

(note- each item on left should be paired with a different one on right)

Weapons-	-far more varied in Late Iron Age
Burials-	-change in materials but not in basic shapes
Food Getting-	-not found in Late Stone Age, but often found in Late Iron Age
Ornaments-	-utilized by groups in Late Stone Age, and by individuals in Late Iron Age
Tools-	-no major change

Please answer some questions about yourself so that we can find out if a representative sample of visitors are answering this questionnaire:

(✓)

A. Is this the first time you have visited this museum?
( )Yes. ( )No- then please estimate the number of times you've visited this museum in the last 12 months _____

B. How interested are you in the early history and archaeology of Scotland?
( )very interested, ( )moderately, ( )slightly, ( )not interested

C. Would you describe your general knowledge of Scottish (or British) archaeology as- ( )very good, ( )good, ( )fair, ( )poor, ( )non-existent

D. Where are you from?
Edinburgh residents- please give district or street-
Others- please give country and town or state-

E. Are you- ( )Male. ( )Female.

F. What is your occupation and what exactly do you do in it-

G. What is your age now- _____

H. At what age did you stop being a full-time student- _____

I. Did you know you were going to be asked to fill-out this questionnaire?
( )No. ( )Yes- then how did you know-

Comments- What did you think of the arrangement of the display in this case?
How would you like to see it improved? - (use back if needed)

Please return questionnaire to box

Thank you very much for your help

Appendix E. Pilot and Control Group Labels:

Appendix E.I. Original Labels in Yellow Case

STONE AGE, ORIGINAL YELLOW LABELS:

The earliest traces of men in Scotland are of Mesolithic (Middle Stone Age) communities who came to Scotland when the ice finally retreated. They lived by hunting, trapping and fishing in small nomadic groups, moving seasonally over traditional areas.

There were two distinct cultures. The one is coastal, the dwelling-places (often caves) are marked by large refuse-heaps mainly consisting of shells. The typical implements are harpoons and limpet scoops. The other, always living on sandy ground, sometimes on the coast, is represented by tiny carefully made flints which were mounted together to form a composite tool. Presumably these people also possessed objects made of wood, bark and leather, but these do not survive.

STONE AGE, YELLOW (NOTE THE DONOR'S NAMES AND THE ACQUISITION DATES):

MESOLITHIC IMPLEMENTS.

Part of a large Secondary Neolithic vessel of coarse ware decorated with impressions of whipped cord. From Glenluce Sands, Wigtownshire. The Executrices of the late J. Smith. 1931.

Cast of harpoon made of deer antler, used for hunting seals and large fish. From the River Dee, Kirkcudbrightshire. Original in Kirkcudbright Museum.

Cores of chert and jasper from which small blades have been struck.

FROM THE VILLAGE AT SKARA BRAE, ORKNEY.

Two views of the interior of the magnificent chambered tomb at Maes Howe, Orkney. (Photographs by the Ancient Monuments Department, Ministry of Works).

STONE AGE, ORIGINAL YELLOW LABELS:

The Neolithic (New Stone Age) is marked by the arrival of new groups of people coming up the Irish Sea from Western Europe, the "Western Neolithic" who were pasturalists and agriculturalists. They made pottery and used new forms of tools. The polished stone axes were an essential piece of equipment to clear the forest, and they were produced in great numbers. The fact that these people had herds and grew grain meant that, unlike the Mesolithic hunters, it was possible to some extent to control their resources so that life was less precarious and the population increased. Most of their relics have been found in chambered cairns, which were family tombs, often of great size and elaborate construction.

Besides these incomers there existed "Secondary Neolithic" peoples, of native mesolithic origin but deeply affected by the new way of life. They made pottery and were extensive traders but do not seem to have adopted agriculture. A complete village belonging to one of the Secondary Neolithic groups was discovered at Skara Brae, Orkney, where it had been inundated by blown sand. The stone huts remained except for the roofs, also the furniture made of stone flags, the tools and the ornaments of the inhabitants.

In the course of time these peoples tended to mix and became less distinct groups - for instance secondary neolithic implements are found in chambered cairns.

Boar's tusk pendant

Small scraper of chert

Stone rubber from Oransay.
J.H. Bishop. 1952

Knappers Farm, Dunbartonshire
FO 965 L. McL. Mann 1956
(P.S.A.S. LXXXI, 234)

A corner of the chamber showing the entrance to one of the subsidiary cells, in the foreground the stone once used to block its entrance.

Leaf-shaped flint arrowheads, the Western Neolithic form, from Unstan, Orkney; Cairnholy, Kirkcudbrightshire; Corriegills, Arran.
Purchased, W. McCulloch and A. Cook, 1885, 1950 and 1909.

arrowheads

Two lop-sided flint arrowheads, a Secondary Neolithic form (compare with the arrowheads below)
From Overhowden Farm, Channelkirk, Berwickshire.
J. Sharp. 1950

Part of a large bowl of Western Neolithic ware, hard dark pottery decorated by faint vertical fluting with the finger-tips. From Easterton of Roseisle, Morayshire.
Mrs H.W. Young. 1921

Polished axe of jadeite (probably imported from Brittany). Found at Comrie Farm, Aberfeldy, Perthshire.

From the chambered cairn at Clettraval, North Uist.
The Trustees of Sir A. Campbell Orde. 1935.

Looking down the entrance passage from the chamber.

Isometric view of house 7, Skara Brae, Orkney.
From the drawing by Professor Piggott.

From the chambered cairn at Unstan, Orkney.
Purchased. 1885.

Knife made from a split pebble.

Polished stone axe found at Biggar Sheils, Lanarkshire. Such axes were quarried at various centres where there is suitable rock, and were traded over most of Britain.

The Bronze Age.

About 1800 B.C. a new people, distinguished by their so-called "Beaker" pottery, came to E.Scotland across the North Sea, and gradually spread over much of the country. About the same time the earliest small bronze objects were brought to Scotland. The use of the new metal developed and the great number of flat axe-heads shows they became general tools replacing the stone implements. In the Middle and Late Bronze Age metal became more abundant and technical knowledge grew so new tools and (for the first time) weapons were produced in increasing numbers.

The only habitation sites of this period yet discovered are some small Late Bronze Age villages in Orkney, where small stone huts with central hearths cluster together. Burials are found fairly frequently in all parts of the country. Beaker-folk generally placed the body doubled up in a stone cist (or box) generally covered by a cairn. Sometimes a pot or other article was placed in the grave. A little later "food-vessels" were the usual offering occasionally with a necklace. Cremation was another custom, which probably continued in use later, the remains being placed in an inverted "cinerary urn", in the ground. These burials seldom have any grave-goods with them.

BRONZE AGE, ORIGINAL YELLOW LABELS:

Crescentic jet necklace found with a Food Vessel, with a short cist burial at Pitkennedy, Angus.
EQ 384 Purchased 1930

Part of a bowl, imported from the Rhineland

Food-vessel, decorated with incised lines and small triangular jabs. From Ratho, Midlothian.

Beaker, decorated with a continuous string-impression. From Aberdeenshire

Bronze halberd - a weapon developed from the bronze dagger but mounted at right angles to a long shaft. No provenance.
Purchased 1887

Bronze dagger, from a cist at Cleigh, Looh Nell, Argyll
Dr.R.A.Smith 1874

Flint dagger, an imitation of the bronze type, found near Nunraw, E. Lothian.
Purchased 1874

Flint arrowheads from Aberdeenshire and Banffshire.
John Smith Collection 1904, 1931

Bronze flat axe, found in a cairn above Upper Towie, Aberdeenshire.
Capt. H.P.Lumsden 1937

Cast of a bronze sickle, found in the Tay near Errol, Perthshire.

Cinerary urn found near Inverkeithing, Fife.
Purchased 1888

Socketed axe, found near Maybole, Ayrshire.
Miss M.G.Macdonald 1951

A cinerary urn as found, placed inverted in a small pit in the ground and enclosed by a setting of stones.

Gold lunula with finely chased decoration.
No provenance.
Purchased 1898

By the last century or two B.C. the Iron Age had begun in Scotland. In the south, at least, Celtic-speaking Britons were dominant, led by tribal warrior-chiefs. In a time of much warfare the warrior-class liked beautiful accoutrements and ornaments, especially bronze-work. Iron was now used freely for better and more plentiful weapons and tools.

Many forts were built for emergencies, but numerous small farmsteads thrived. A few fortified hill-top towns were established as tribal centres, such as that on Traprain Law. In the north, particularly, brochs (stone towers, really defensive farmhouses) were built in great numbers.

The Romans were led to protect their Empire by conquering southern Britain and then, about 80 A.D. advancing into Scotland. The south of Scotland lay within the Empire intermittently until at least 185 A.D. but the north always remained free. Life was not fundamentally affected, but the native Celts obtained many Roman articles and adopted some Roman techniques.

IRON AGE, ORIGINAL YELLOW LABELS:

Stone whorl, to weight the end of a spindle for spinning yarn. From the Broch of Burrian, N. Ronaldsay, Orkney.
Dr W. Traill. 1872

Wheel-made Roman vessel. Roman pottery was used extensively by the native population.
From Traprain Law

Tip of a sword. From Carlingwark Loch, Kirkcudbrightshire.
W.W. Hay Newton. 1852.

Butt of a spear-shaft, bronze with an iron tip.
From Traprain Law

All from Traprain Law, most of them made in the village.

Shears

Pieces of glass and jet armlets, from Traprain Law.

Yellow glass beads, from the Culbin Sands, Morayshire.

Isometric reconstruction of an Iron Age crannog house (Milton Loch, Kirkcudbrightshire, from the drawing by Mrs C.M. Piggott)

Bronze armlet, one of a pair found on the Links of Drumside, Belhelvie, Aberdeenshire.
Dr J. Stuart. 1853.

Stone ball, perhaps used as a sling-stone.

Bronze fibula with open-work catchplate. From Newpart Farm, Polmaise, Stirlingshire.
L.1956.4. Lt. Commander G.A. Hamilton.

Jar of Roman manufacture.
From Traprain Law.

Reconstructed view of the Iron Age settlement at Jarlshof, Shetland. The broch is no longer used and has been partially dismantled, five stone-built circular huts are now used as dwellings. From the drawing by Alan Sorrell.

PERSONAL ORNAMENTS

Another North British form.
2nd century A.D.

Bronze terret, for holding the reins on a war chariot, such as were extensively used by Iron Age warriors. From Traprain Law.

Cast of an iron sickle

Small penannular brooch.

POST ROMAN, ORIGINAL YELLOW LABELS:

In the Post-Roman period three Celtic kingdoms were established in Scotland: that of the Picts who emerged as a nation from the Iron Age tribes of E. & NE. Scotland, that of the Britons of Strathclyde in SW. Scotland, and that of the Scots who were incomers to Argyll from Ireland. During much of the period south Scotland was subject to the powerful Anglian kingdom of Northumbria. From about 800 A.D. the Shetland and Orkney Islands were colonised by Norwegians, and the Vikings made raids and eventually settled in NE. Scotland and the Hebrides.

South Scotland was converted to Christianity towards the end of the Roman Empire, and much of the rest of Scotland as a result of the mission of St Columba in the 6th century. Iona (in the Scottic kingdom) became the great religious centre, and the church must have been a unifying influence before the eventual uniting of the kingdoms under the king of the Scots.

The archaeological material of the period is meagre and few sites have been excavated, two exceptions being the Viking settlement at Jarlshof, Shetland and the Scottic fortress at Dunadd, Argyll. Celtic craftsmen made superb metal work, especially brooches, and there were also beautiful carved stone monuments made by the Picts.

POST ROMAN, ORIGINAL YELLOW LABELS:

CELTIC PENANNULAR BROOCHES

Knives.

Photographs of two standing stones with Pictish carving:
1. A battle scene, with symbols above (at Aberlemno, Angus)
2. With symbols only, at the top a 'flower-symbol', in the centre a double-circle and Z-rod, at the base a comb and mirror; it is not known what the symbols mean (at Dunnichen, Angus).
Other Pictish stones are exhibited on the ground floor of the museum.

Parts of clay moulds for casting Celtic penannular brooches with interlace design, with wax impressions taken from them. From the fort of Mote of Mark, Dalbeattie, Kirkcudbrightshire.
Governors of the Hutton Trust. 1913

FC 177. Penannular brooch of silver, th
expanded ends inlaid with decorative gol
plates. About 780 A.D. Found nea
Perth.
Purchased. 190

Bone dice. Found at St Tredwell's Chapel, Papa Westray, Orkney.
Dr W. Traill. 1882

Bone comb and pins from the broch of Burrian, N.Ronaldsay, bronze pin and needle from the broch of Lamaness, Sanday.
Dr W. Traill. 1872
Col. D. Balfour. 1878

Iron axe-head, bronze belt-buckle and bronze pin, found with a sword and shield boss in a Viking's grave at Ballinaby, Islay, Argyll.
A. McLachlan. 1933.

A GROUP OF OBJECTS FROM THE ORKNEY ISLANDS

Silver plaque with Pictish symbols such as are found on the carved stone monuments (see below). (The red inlay is a modern replacement). From the hoard of silver ornaments found at Norrie's Law, Largo, Fife.
R. Dundas. 1883

Appendix E.II. Labels in the Pilot Display Case

MESOLITHIC 6,000 - 3,500 BC

Nothing is left to indicate the appearance of Scotlands' first human inhabitants, what clothes they wore, or how they were buried. Although some of their remains have been found in caves, no one can be sure of what their other dwellings looked like. Information about them may however be inferred from archaeological discoveries about contemporary Middle Stone Age groups; from biological and geological, analysis; and from knowledge of similar-appearing cultures today, such as that of the Caribou Eskimos.

People first ventured into Scotland sometime before 6,000 BC. The earliest traces yet found are of Middle Stone Age (Mesolithic) hunters and gatherers who settled here after the last ice-sheets had retreated. These small nomadic groups are recognised in the archaeological record primarily from coastal shell heaps containing the remains of tools and bones. Since they had to carry all of their goods with them when they moved in search of food, they probably didn't possess much in the way of large or easily breakable goods such as pottery. Beads from their shell necklaces indicate that some interest was paid to individual ornaments.

MESOLITHIC IMPLEMENTS

Cast of harpoon made of deer antler, used for hunting seals and large fish. From the River Dee, Kirkcudbrightshire.

Original in Kirkcudbright Museum.

Tiny flints of various forms, used by mounting a number in a row in a wooden shaft.

Limpet 'hammer', probably used to detach limpets from their rocks.

Cowrie shell necklace, an illustration of how Mesolithic jewellery might have appeared.

A Mesolithic burial of a woman with a new-born baby at Bøgebakken, Vedbaek, Denmark. The baby was laid on a swan's wing and there is a larger number of pendants of pig and deer teeth near the woman's head.

NEOLITHIC 3,500 - 2,000 BC

Although hunting and gathering continued with decreasing importance until Viking times, the basic agricultural practices which mark the beginning of the Late Stone Age (Neolithic) were to remain unchanged until the Industrial Revolution over 20 centuries later. With a more secure and efficient foodbase in agriculture, more permanent settlements are now found and a population increase seems a reasonable assumption. An increase in the variety of material goods is also witnessed by the introduction of pottery. Axes come into prominence to clear the land for crops and animals.

Most of our information about these Neolithic farmers comes from their burials and from a few of their settlements. None of their settlements appear to have been defensive and no identifiable weapons have been found. The new social structures which gradually must have developed from their stable base are recognisable in the archaeological record by the appearance of large tombs for collective burials and, later, by large ritual structures such as stone circles, which presuppose an economic surplus. Evidence of agriculture in the prehistoric record takes two forms:

i finds of grain (carbonised) or domestic animal bones
ii finds of agricultural equipment.

Leaf shaped flint arrowheads, the Western Neolithic form, from Unstan, Orkney; Cairnholy, Kirkcudbrightshire; Corriegills, Arran.

Purchased, W. McCullock and A. Cook, 1885, 1950 and 1909.

Saddle quern, used to grind grain.

Bone awl and flint scraper, imply leather was being worked.

Small scraper of chert.

bone awl

Polished stone axe found at Biggar Shiels, Lanarkshire. Such axes were quarried at various centres where there is suitable rock, and were traded over most of Britain.

Polished axe of jadeite (probably imported from Brittany). Found at Comrie Farm, Aberfeldy, Perthshire.

BRONZE AGE 2000 - 600 BC

Metalworking was brought to Scotland around 2000 BC by Beaker groups from the Low Countries and north Germany. In addition to bronzework Beaker people also introduced a distinctive pottery style (from which they are named) and the use of individual crouched burials. Since the ability to make bronze doesn't appear to have been widespread until the end of the period, stone tools remained important. Definite fighting implements first occur in the Bronze Age. They could have been used more as marks of personal prestige than for serious fighting. Large numbers of weapons and the appearance of the first recognisable defensive settlements (hill forts) indicate the presence of warefare in the Late Bronze Age.

Early burials of crouched individuals inside stone cists (or boxes) often contain Beaker pots. These pots are later joined, then replaced, by Food Vessel pottery. Even before the end of the Early Bronze Age, however, cremation in inverted cinerary urns largely replaces the use of individual graves. Later in the Bronze Age the absence of grave goods prevents the recognition of differing traditions. Settlements of this period are rare. Although burials provide much information until 1400 BC, thereafter we are largely dependent on hoards of metal objects to interpret developments. Increasing technical skill is evident in objects intended for personal ornament as well as for weapons. Luxury materials such as gold and jet also come into use.

Cast of a bronze sickle, found in the Tay, near Errol, Perthshire.

Socketed gouge. No provenance.
Sim Collection.

Socketed chisel. No provenance.
Sim Collection.

Knife.

Socketed axe, found near Maybole, Aryshire.
Miss M. G. Macdonald, 1951.

Flint dagger, an imitation of the bronze type, found near Nunraw, East Lothian.

EARLY IRON AGE 600 BC - 300 AD

Iron was probably brought to Scotland by Celtic warriors. Since the ore was relatively easy to find it must have allowed better and more plentiful tools to be used in all aspects of life. Bronze jewellery seems to have been preferred throughout the period. The ability to cast iron was not known at this time so finer decorative metalwork was possible only in bronze.

Iron objects don't survive well in the archaeological record. Consequently iron implements in Scotland are primarily known from a few scrap hoards, two of which were found in bronze cauldrons. Numerous settlements are known from the Early Iron Age but, except for those in the north and west, only occasionally produce many objects when excavated. Large defensive hill forts and brochs (tower-like farmhouses) remain as well as evidence of many small farmsteads and forts. Warfare is implied by the large number of defensive settlements in the Early Iron Age and by the occasional find of weapons. Classical writers also emphasised the importance of battle skills. Examples of Early Iron Age burials are rare. The few which have been dated to this period contained skeletons with few grave goods.

Cast of an iron sickle.

Shears.

Ox goad.

Sock of a plough.

Iron Age ploughs were made of wood. By fitting an iron plough sock to the underside of the share a considerable reduction in wear was achieved.

Bone weaving comb, for pushing the threads into place during weaving. From Burgar, Orkney.

Mr Peterkin, 1827.

Clay loomweight.

Whorl and reconstructed spindle used to spin wool.

Spindle-whorl made of a Roman pot-sherd.

All imply use of cloth in the Early Iron Age.

LATE IRON AGE 300 - 1100 AD

War fare continues in the Late Iron Age (Early Christian Period) well after the Romans have left Britain, late in the fourth century AD. It is known from historical sources that at least three kingdoms were formed in Scotland - that of the Picts, the Britons, and the Scots (incomers to Argyll from Ireland). Later in the period, around 800 AD, Vikings first plundered in Scotland then settled large areas in the north and west. Most of the richest Viking material has been found in hoards or burials. The latter, with their rich grave goods, contrasted markedly with the plain unaccompanied burials of the Christianised Late Iron Age communities over much of Scotland.

Other Archaeological material of the period is meagre and few sites have been excavated apart from the Viking settlement at Jarlshof, Shetland and the Scottic fortress at Dunadd, Argyll. Because utilitarian objects remained similar throughout the Iron Age, chronological divisions are difficult and are only really possible in the changing field of art-styles.

PICTISH SILVER

Silver plaque with Pictish symbols such as are found on the carved stone monuments. (The red inlay is a modern replacement). From the hoard of silver ornaments found at Norrie's Law, Largo, Fife.

R. Dundas, 1883.

Silver pin, the head originally inlaid, decorated with a scroll and triquetra motif, a small Maltese cross as above. On the back is engraved the Pictish "V-rod" symbol. From the Norrie's Law Hoard.

FC 177. Penannular brooch of silver, the expanded ends inlaid with decorative gold plates. About 780 A.D. Found near Perth.

Purchased, 1906.

Bead of yellow, white and blue glass.

Bronze and bone pins.

Bone dice.

Bronze needle.

Bone combs.

Appendix F. Alleval Study:

Appendix F.I. Alleval Questionnaire (reduced):

Introductory Case Evaluation

please fill-in the following questions about yourself so that we can find out if a representative sample of visitors are helping us evaluate the display case:

sex - _Male _Female

age - _____

occupation -

residence -

date: 7 June 1979

sheet no:

	RATING				
	VERY POOR				EXCELLENT
	1	2	3	4	5
note - these are general questions requiring no specific technical skill					
1. Is the purpose of the display case clearly communicated to you?					
2. How effective is the display in raising questions about its content?					
3. How well does the display serve the casual visitor?					
4. How well does the display serve the scholar or expert?					
5. Are the start and finish of the case clearly indicated or apparent?					
6. Is it easy to see all the objects in the case that you wish to see?					
7. Is the size of print (type) easily read in the cream labels? — .... in the green labels? —					
8. Does the arrangement of the display maintain your interest throughout?					
9. Does the lighting of the case enhance the artifacts?					
10. Is the display case well maintained?					
11. Are security arrangements adequately concealed?					
12. How effective is this display case in making you wish to return for another visit?					

your remarks:

After carefully studying the Introductory Case please tick (✓) only those categories below which you strongly feel should be changed.
exception - In the two starred categories (★) please indicate whether you prefer item a. or b. in each.

IN this display case I would like to see:

_ more general information
_ less general information

_ more objects
_ fewer objects

_ easier vocabulary
_ more technical vocabulary

_ more information about where these objects were found
_ more information about how archaeologists make their discoveries
_ maps for reference to finds

_ the arrangement changed so that it all reads from left to right
ie- start with STONE age on left-hand side of case

★ a. _ major information broken-up beside each group of objects as in the Bronze Age and Early Iron Age displays.
★ b. _ major information put on one large label as in the Stone Age and Late Iron Age displays.

★ a. _ the artifacts grouped into themes of tools, weapons, etc where possi
★ b _ no particular grouping of the artifacts (as in Late Iron Age)

_ more comparisons drawn to contemporary cultures

I would welcome any further comments or suggestions -

Appendix F.II. Alleval Study Results:

SEX

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
MALE	1.	10	32.3	40.0	40.0
FEMALE	2.	15	48.4	60.0	100.0
no answer	0.	6	19.4	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.100	MEDIAN	1.667	MODE	2.000
VARIANCE	0.250	RANGE	1.000		
VALID CASES	25	MISSING CASES	6		

AGE

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
18-29	2.	16	51.6	61.5	61.5
30-49	3.	8	25.8	30.8	92.3
50-83	4.	2	6.5	7.7	100.0
no answer	0.	5	16.1	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.127	MEDIAN	2.312	MODE	2.000
VARIANCE	0.418	RANGE	2.000		
VALID CASES	26	MISSING CASES	5		

WORK SOCIAL ECONOMIC GROUP JOB NUMBER

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
PROFESSIONALS	1.	12	38.7	46.2	46.2
OFFICE WORKERS	2.	4	12.9	15.4	61.5
TRADES&SERVICES	3.	6	19.4	23.1	84.6
STUDENTS	9.	4	12.9	15.4	100.0
no answer	0.	5	16.1	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.548	MEDIAN	1.750	MODE	1.000
VARIANCE	7.815	RANGE	8.000		
VALID CASES	26	MISSING CASES	5		

RESIDE RESIDENCE

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
EDINBURGH	10.	5	16.1	20.8	20.8
SCOTLAND	20.	3	9.7	12.5	33.3
ENGLAND	30.	2	6.5	8.3	41.7
EUROPE	50.	6	19.4	25.0	66.7
USA&CANADA	69.	8	25.8	33.3	100.0
no answer	0.	7	22.6	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	4.866	MEDIAN	48.333	MODE	69.000
VARIANCE	568.166	RANGE	59.000		
VALID CASES	24	MISSING CASES	7		

RAT1 **IS THE PURPOSE OF THE DISPLAY CASE CLEAR?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
POOR	2.	3	9.7	10.3	10.3
AVERAGE	3.	6	19.4	20.7	31.0
GOOD	4.	10	32.3	34.5	65.5
EXCELLENT	5.	10	32.3	34.5	100.0
no answer	0.	2	6.5	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.185	MEDIAN	4.050	MODE	4.000
VARIANCE	0.995	RANGE	3.000		
VALID CASES	29	MISSING CASES	2		

RAT2 **HOW EFFECTIVE IN RAISING QUESTIONS?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	1	3.2	3.6	3.6
POOR	2.	4	12.9	14.3	17.9
AVERAGE	3.	5	16.1	17.9	35.7
GOOD	4.	13	41.9	46.4	82.1
EXCELLENT	5.	4	12.9	14.3	96.4
other	8.	1	3.2	3.6	100.0
no answer	0.	3	9.7	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.251	MEDIAN	3.808	MODE	4.000
VARIANCE	1.767	RANGE	7.000		
VALID CASES	28	MISSING CASES	3		

RAT3 **HOW WELL SERVE THE CASUAL VISITORS?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	1	3.2	3.4	3.4
POOR	2.	4	12.9	13.8	17.2
AVERAGE	3.	8	25.8	27.6	44.8
GOOD	4.	8	25.8	27.6	72.4
EXCELLENT	5.	8	25.8	27.6	100.0
no answer	0.	2	6.5	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.213	MEDIAN	3.687	MODE	3.000
VARIANCE	1.315	RANGE	4.000		
VALID CASES	29	MISSING CASES	2		

RAT4 **HOW WELL SERVE THE SCHOLAR OR EXPERT?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	2	6.5	7.7	7.7
POOR	2.	4	12.9	15.4	23.1
AVERAGE	3.	8	25.8	30.8	53.8
GOOD	4.	7	22.6	26.9	80.8
EXCELLENT	5.	1	3.2	3.8	84.6
other (see COM6)	6.	4	12.9	15.3	100.0
no answer	0.	5	16.1	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.319	MEDIAN	3.375	MODE	3.000
VARIANCE	2.654	RANGE	7.000		
VALID CASES	26	MISSING CASES	5		

RAT5 **ARE START AND FINISH CLEARLY INDICATED?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	1	3.2	3.4	3.4
POOR	2.	6	19.4	20.7	24.1
AVERAGE	3.	3	9.7	10.3	34.5
GOOD	4.	13	41.9	44.8	79.3
EXCELLENT	5.	6	19.4	20.7	100.0
no answer	0.	2	6.5	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.214	MEDIAN	3.846	MODE	4.000
VARIANCE	1.323	RANGE	4.000		
VALID CASES	29	MISSING CASES	2		

RAT6 **IS IT EASY TO SEE OBJECTS IN THIS DISPLAY CASE?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
POOR	2.	4	12.9	14.8	14.8
AVERAGE	3.	5	16.1	18.5	33.3
GOOD	4.	6	19.4	22.2	55.6
EXCELLENT	5.	12	38.7	44.4	100.0
no answer	0.	4	12.9	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.217	MEDIAN	4.250	MODE	5.000
VARIANCE	1.268	RANGE	3.000		
VALID CASES	27	MISSING CASES	4		

RAT7A **EASY TO READ TYPESIZE IN CREAM LABELS?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	2	6.5	7.4	7.4
POOR	2.	1	3.2	3.7	11.1
AVERAGE	3.	4	12.9	14.8	25.9
GOOD	4.	7	22.6	25.9	51.9
EXCELLENT	5.	13	41.9	48.1	100.0
no answer	0.	4	12.9	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.236	MEDIAN	4.429	MODE	5.000
VARIANCE	1.499	RANGE	4.000		
VALID CASES	27	MISSING CASES	4		

RAT7B **EASY TO READ TYPESIZE IN GREEN LABELS?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	4	12.9	16.0	16.0
POOR	2.	5	16.1	20.0	36.0
AVERAGE	3.	7	22.6	28.0	64.0
GOOD	4.	6	19.4	24.0	88.0
EXCELLENT	5.	3	9.7	12.0	100.0
no answer	0.	6	19.4	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.255	MEDIAN	3.000	MODE	3.000
VARIANCE	1.623	RANGE	4.000		
VALID CASES	25	MISSING CASES	6		

RAT8 **DOES DISPLAY ARRANGEMENT KEEP YOUR INTEREST?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	1	3.2	3.6	3.6
POOR	2.	2	6.5	7.1	10.7
AVERAGE	3.	6	19.4	21.4	32.1
GOOD	4.	12	38.7	42.9	75.0
EXCELLENT	5.	7	22.6	25.0	100.0
no answer	0.	3	9.7	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.195	MEDIAN	3.917	MODE	4.000
VARIANCE	1.063	RANGE	4.000		
VALID CASES	28	MISSING CASES	3		

RAT9 **DOES LIGHTING ENHANCE THE ARTIFACTS?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
VERY POOR	1.	5	16.1	17.9	17.9
POOR	2.	5	16.1	17.9	35.7
AVERAGE	3.	6	19.4	21.4	57.1
GOOD	4.	9	29.0	32.1	89.3
EXCELLENT	5.	2	6.5	7.1	96.4
other	6.	1	3.2	3.6	100.0
no answer	0.	3	9.7	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.260	MEDIAN	3.167	MODE	4.000
VARIANCE	1.888	RANGE	5.000		
VALID CASES	28	MISSING CASES	3		

RAT10 **IS THE DISPLAY CASE WELL MAINTAINED?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
AVERAGE	3.	3	9.7	10.7	10.7
GOOD	4.	12	38.7	42.9	53.6
EXCELLENT	5.	13	41.9	46.4	100.0
no answer	0.	3	9.7	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.128	MEDIAN	4.417	MODE	5.000
VARIANCE	0.460	RANGE	2.000		
VALID CASES	28	MISSING CASES	3		

RAT11 **ARE SECURITY ARRANGEMENTS CONCEALED?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
AVERAGE	3.	1	3.2	3.8	3.8
GOOD	4.	7	22.6	26.9	30.8
EXCELLENT	5.	17	54.8	65.4	96.2
other	6.	1	3.2	3.8	100.0
no answer	0.	5	16.1	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.121	MEDIAN	4.794	MODE	5.000
VARIANCE	0.382	RANGE	3.000		
VALID CASES	26	MISSING CASES	5		

RAT12 **DOES CASE MAKE YOU WISH TO REVISIT MUSEUM?**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
POOR	2.	3	9.7	10.3	10.3
AVERAGE	3.	14	45.2	48.3	58.6
GOOD	4.	5	16.1	17.2	75.9
EXCELLENT	5.	7	22.6	24.1	100.0
no answer	0.	2	6.5	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.183	MEDIAN	3.321	MODE	3.000
VARIANCE	0.970	RANGE	3.000		
VALID CASES	29	MISSING CASES	2		

QINFO **GENERAL INFORMATION WANTED:**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
MORE	1.	11	35.5	91.7	91.7
LESS	2.	1	3.2	8.3	100.0
no comment	0.	19	61.3	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.083	MEDIAN	1.045	MODE	1.000
VARIANCE	0.083	RANGE	1.000		
VALID CASES	12	MISSING CASES	19		

QOBJ **OBJECTS WANTED:**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
MORE	1.	5	16.1	31.2	31.2
FEWER	2.	11	35.5	68.7	100.0
no comment	0.	15	48.4	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.120	MEDIAN	1.773	MODE	2.000
VARIANCE	0.229	RANGE	1.000		
VALID CASES	16	MISSING CASES	15		

QVOCAB **VOCABULARY WANTED:**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
EASIER	1.	4	12.9	57.1	57.1
MORE TECHNICAL	2.	3	9.7	42.9	100.0
no comment	0.	24	77.4	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.202	MEDIAN	1.375	MODE	1.000
VARIANCE	0.286	RANGE	1.000		
VALID CASES	7	MISSING CASES	24		

QMAPS **MORE INFORMATION ABOUT-**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
WHERE OBJ FOUND	1.	6	19.4	33.3	33.3
HOW ARCH. FIND	2.	1	3.2	5.6	38.9
MAP FOR FINDS	3.	9	29.0	50.0	88.9
BOTH WHERE & HOW (ABOVE)	4.	1	3.2	5.6	94.4
BOTH WHERE & MAP (ABOVE)	5.	1	3.2	5.6	100.0
no comment	0.	13	41.9	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.283	MEDIAN	2.722	MODE	3.000
VARIANCE	1.438	RANGE	4.000		
VALID CASES	18	MISSING CASES	13		

QRTLF **CHANGE TO READ LEFT TO RIGHT:**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
YES	1.	11	35.5	100.0	100.0
no comment	0.	20	64.5	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.000	MEDIAN	1.000	MODE	1.000
VARIANCE	0.000	RANGE	0.000		
VALID CASES	11	MISSING CASES	20		

QLABELS **MAJOR INFORMATION WANTED:**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
BROKENUP BY OBJ	1.	17	54.8	70.8	70.8
ON ONE LARGE LABEL	2.	7	22.6	29.2	100.0
no comment	0.	7	22.6	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.095	MEDIAN	1.206	MODE	1.000
VARIANCE	0.216	RANGE	1.000		
VALID CASES	24	MISSING CASES	7		

QARTFCT **ARTIFACT GROUPING PREFERRED:**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
IN THEMES (T.W.O.B.)	1.	18	58.1	85.7	85.7
NO GROUPING, UNSTRUCTURED	2.	3	9.7	14.3	100.0
no comment	0.	10	32.3	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.078	MEDIAN	1.083	MODE	1.000
VARIANCE	0.129	RANGE	1.000		
VALID CASES	21	MISSING CASES	10		

QCONTEMP **MORE COMPARISONS TO CONTEMPORARY CULTURE:**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
YES, WANT MORE	1.	5	16.1	100.0	100.0
no answer	0.	26	83.9	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.000	MEDIAN	1.000	MODE	1.000
VARIANCE	0.000	RANGE	0.000		
VALID CASES	5	MISSING CASES	26		

COM1 **DIFFICULTY READING LABELS BECAUSE OF-**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
REFL.& BG COLOUR	21.	2	6.5	33.3	33.3
PRINT SIZE & PLACEMENT	22.	3	9.7	50.0	83.3
BOTH OF ABOVE	215.	1	3.2	16.7	100.0
null	0.	25	80.6	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	32.234	MEDIAN	21.833	MODE	22.000
VARIANCE	6234.164	RANGE	194.000		
VALID CASES	6	MISSING CASES	25		

COM2 **LABEL IMPROVEMENTS SUGGESTED**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
MORE GENERAL INFORMATION	31.	2	6.5	33.3	33.3
LESS INFORMATION	32.	4	12.9	66.7	100.0
null	0.	25	80.6	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.211	MEDIAN	31.750	MODE	32.000
VARIANCE	0.267	RANGE	1.000		
VALID CASES	6	MISSING CASES	25		

COM3 **CRITICAL OF THE DISPLAY**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
DISLIKE, SUGG.IMPROVEMENTS	32.	4	12.9	57.1	57.1
NOTE FAULTS, SUGG.IMPROV.	33.	3	9.7	42.9	100.0
null	0.	24	77.4	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.202	MEDIAN	32.375	MODE	32.000
VARIANCE	0.286	RANGE	1.000		
VALID CASES	7	MISSING CASES	24		

COM4 **EMPHASIS ON OBJECTS**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
OBJECT CENTRED	10.	2	6.5	100.0	100.0
null	0.	29	93.5	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.000	MEDIAN	10.000	MODE	10.000
VARIANCE	0.000	RANGE	0.000		
VALID CASES	2	MISSING CASES	29		

COM5 **CRITICAL OF THE THEMATIC ARRANGEMENT**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
TOO FRAGMENTED	12.	2	6.5	100.0	100.0
null	0.	29	93.5	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.000	MEDIAN	12.000	MODE	12.000
VARIANCE	0.000	RANGE	0.000		
VALID CASES	2	MISSING CASES	29		

COM6 **NOT EXPERT SO CAN'T ANSWER QUESTION 4**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
DENY BEING EXPERT	21.	3	9.7	60.0	60.0
EXPERT NEEDS-MORE OR LESS	22.	2	6.5	40.0	100.0
null	0.	26	83.9	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.245	MEDIAN	21.333	MODE	21.000
VARIANCE	0.300	RANGE	1.000		
VALID CASES	5	MISSING CASES	26		

COM7 **SOCIAL CONTEXT EMPHASIZED**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
GENERAL HUMAN INTEREST	21.	2	6.5	50.0	50.0
SPECIFIC HUMAN INTEREST	22.	1	3.2	25.0	75.0
BOTH OF THE ABOVE	34.	1	3.2	25.0	100.0
null	0.	27	87.1	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	3.175	MEDIAN	21.500	MODE	21.000
VARIANCE	40.333	RANGE	13.000		
VALID CASES	4	MISSING CASES	27		

COM8 **EXCUSES**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
TIME OR INABILITY EXCUSE	10.	5	16.1	100.0	100.0
null	0.	26	83.9	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.000	MEDIAN	10.000	MODE	10.000
VARIANCE	0.000	RANGE	0.000		
VALID CASES	5	MISSING CASES	26		

COM9 **THEMATIC FORMAT PREFERRED**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
CONTINUOUS FLOW	31.	2	6.5	25.0	25.0
THEMES EMPHASIZED	32.	5	16.1	62.5	87.5
LEVELS OF DISPLAY SUGG.	33.	1	3.2	12.5	100.0
null	0.	23	74.2	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.227	MEDIAN	31.900	MODE	32.000
VARIANCE	0.411	RANGE	2.000		
VALID CASES	8	MISSING CASES	23		

SHEETORD **PAGE ORDER 1-2 OR 2-1**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
ONE-TWO	1.	24	77.4	77.4	77.4
TWO-ONE	2.	7	22.6	22.6	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.076	MEDIAN	1.146	MODE	1.000
VARIANCE	0.181	RANGE	1.000		
VALID CASES	31	MISSING CASES	0		

COM10 **GENERAL APPROVAL OF DISPLAY**

	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
LIKE	21.	2	6.5	50.0	50.0
LIKE UNDER THE CIRCUMSTANCES	22.	2	6.5	50.0	100.0
null	0.	27	87.1	MISSING	100.0
	TOTAL	31	100.0	100.0	

STD ERR	0.289	MEDIAN	21.500	MODE	21.000
VARIANCE	0.333	RANGE	1.000		
VALID CASES	4	MISSING CASES	27		

Appendix G. Control Study Results:

PILOT QUESTIONNAIRE AND PRE-TEST= DISPLAY UNSEEN;
POST-TEST, ASK (TO) STUDY, TEST (YOUR) SELF, AND FIRHILL HIGH SCHOOL GROUP= SAW DISPLAY.

Appendix G.I. Test Score by Control Experiments

C R O S S T A B U L A T I O N O F
TCORECT- TOTAL CORRECT TEST SCORE
BY CASEVAR- CONTROL STUDY VARIATION TESTED

	CASEVAR					
COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
		pre*post				
TCORECT						
0.	2	3 (4) 1	0	0	0	6
	3.4	14. 5.	0.0	0.0	0.0	2.8
1-3.	26	8(12) 4	5	2	4	49
	44.1	36. 18.	6.6	16.7	20.0	23.2
4-6.	25	11(15) 4	7	3	2	52
	42.4	50. 18.	9.2	25.0	10.0	24.6
7-9.	4	0 (8) 8	15	3	2	32
	6.8	0. 36.	19.7	25.0	10.0	15.2
10-18.	2	0 (5) 5	49	4	12	72
	3.4	0. 23.	64.5	33.3	60.0	34.1
COLUMN TOTAL	59	22(44)22	76	12	20	211
	28.0	20.9	36.0	5.7	9.5	100.0

12 OUT OF 25 (48.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.341
CHI SQUARE = 100.72154 WITH 16 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

Appendix G.II. First Visit by Control Treatments

C R O S S T A B U L A T I O N O F
FIVISIT- A1) IS THIS YOUR FIRST VISIT TO THIS MUSEUM?
BY CASEVAR- CONTROL STUDY VARIATION TESTED

	CASEVAR					
COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
FIVISIT						
YES	34	21	61	10	16	142
	60.7	95.5	82.4	83.3	88.9	78.0
NO	22	1	13	2	2	40
	39.3	4.5	17.6	16.7	11.1	22.0
COLUMN TOTAL	56	22	74	12	18	182
	30.8	12.1	40.6	6.6	9.9	100.0

2 OUT OF 10 (20.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 2.513
CHI SQUARE = 18.95010 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0008

NUMBER OF MISSING OBSERVATIONS = 7

Appendix G.III. Number of Visits in Past Year by Control Treatments

C R O S S T A B U L A T I O N O F
TIMESYR- A2) NO.VISITS IN PAST YEAR
BY CASEVAR- CONTROL STUDY VARIATION TESTED

TIMESYR	CASEVAR: PILOT UNSEEN	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
1-2.	13	5	1	1	20
	86.7	71.4	100.0	33.3	76.9
3-6.	1	2	0	2	5
	6.7	28.6	0.0	66.7	19.2
12.	1	0	0	0	1
	6.7	0.0	0.0	0.0	3.8
COLUMN TOTAL	15	7	1	3	26
	57.7	26.9	3.8	11.5	100.0

10 OUT OF 12 (83.3%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.038
CHI SQUARE = 7.00761 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.3201

NUMBER OF MISSING OBSERVATIONS = 185

Appendix G.IV. Interest in Scottish History by Control Treatments

C R O S S T A B U L A T I O N O F
INTRST- B) INTEREST IN EARLY SCOTTISH HISTORY
BY CASEVAR- CONTROL STUDY VARIATION TESTED

INTRST	CASEVAR COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
VERY INTERESTED		8	2	12	2	4	28
		15.1	9.1	16.2	16.7	22.2	15.6
MODERATELY		20	9	40	2	9	80
		37.7	40.9	54.1	16.7	50.0	44.7
SLIGHTLY-NOT INT		25	11	22	8	5	71
		47.2	50.0	29.7	66.7	27.8	39.7
COLUMN TOTAL		53	22	74	12	18	179
		29.6	12.3	41.3	6.7	10.1	100.0

4 OUT OF 15 (26.7%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 1.851
CHI SQUARE = 13.08915 WITH 8 DEGREES OF FREEDOM
SIGNIFICANCE = 0.1088

NUMBER OF MISSING OBSERVATIONS = 10

Appendix G.V. Knowledge of Scottish Archaeology by Control Treatments

C R O S S T A B U L A T I O N O F
KNOWLDG- C) AVOWED GENERAL KNOWLEDGE OF SCOTTISH ARCHAEOLOGY
BY CASEVAR- CONTROL STUDY VARIATION TESTED

KNOWLDG	COUNT / COL PCT	CASEVAR: PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
FAIR-V.GOOD KNOW		14	5	31	8	11	69
		26.4	22.7	41.3	66.7	61.1	38.3
POOR-NO KNOWLEDG		39	17	44	4	7	111
		73.6	77.3	58.7	33.3	38.9	61.7
COLUMN TOTAL		53	22	75	12	18	180
		29.4	12.2	41.7	6.7	10.0	100.0

1 OUT OF 10 (10.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 4.508
CHI SQUARE = 14.81918 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0051

NUMBER OF MISSING OBSERVATIONS = 9

Appendix G.VI. Residence by Control Treatments

C R O S S T A B U L A T I O N O F
RESIDENC- D) RESIDENCE
BY CASEVAR- CONTROL STUDY VARIATION TESTED

RESIDENC	COUNT / COL PCT	CASEVAR: PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
EDINBURGH		21	7	16	12	2	58
		38.9	31.8	22.2	100.0	12.5	33.0
SCOTLAND		12	4	8	0	4	28
		22.2	18.2	11.1	0.0	25.0	15.9
ENGLAND		16	4	13	0	4	37
		29.6	18.2	18.1	0.0	25.0	21.0
U.S.A		5	4	27	0	6	42
		9.3	18.2	37.5	0.0	37.5	23.9
UNSPECIFIED ALIEN		0	3	8	0	0	11
		0.0	13.6	11.1	0.0	0.0	6.2
COLUMN TOTAL		54	22	72	12	16	176
		30.7	12.5	40.9	6.8	9.1	100.0

12 OUT OF 25 (48.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.913
CHI SQUARE = 61.01791 WITH 16 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 13

Appendix G.VII. Sex by Control Treatments

C R O S S T A B U L A T I O N O F
SEX- E) SEX
BY CASEVAR- CONTROL STUDY VARIATION TESTED

	CASEVAR					
COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
SEX						
MALE	30 / 57.7	9 / 40.9	36 / 48.0	12 / 100.0	9 / 50.0	96 / 53.6
FEMALE	22 / 42.3	13 / 59.1	39 / 52.0	0 / 0.0	9 / 50.0	83 / 46.4
COLUMN TOTAL	52 / 29.0	22 / 12.3	75 / 41.9	12 / 6.7	18 / 10.1	179 / 100.0

CHI SQUARE = 13.95198 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0074

NUMBER OF MISSING OBSERVATIONS = 10

Appendix G.VIII. Occupation by Control Treatments

C R O S S T A B U L A T I O N O F
OCCSEG F1) OCCUPATION-SOCIAL ECONOMIC GROUP
BY CASEVAR- CONTROL STUDY VARIATION TESTED

	CASEVAR					
COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
OCCSEG						
PROFESSIONAL	9 / 18.7	4 / 19.0	10 / 13.9	0 / 0.0	3 / 17.6	26 / 15.3
MANAGERS	9 / 18.7	3 / 14.3	22 / 30.6	0 / 0.0	5 / 29.4	39 / 22.9
SKILLED	6 / 12.5	3 / 14.3	8 / 11.1	0 / 0.0	3 / 17.6	20 / 11.8
PARTLY SKILLED	2 / 4.2	0 / 0.0	0 / 0.0	0 / 0.0	0 / 0.0	2 / 1.2
STUDENT	12 / 25.0	8 / 38.1	23 / 31.9	12 / 100.0	5 / 29.4	60 / 35.3
RET,UNEMP,TOURIS	7 / 14.6	0 / 0.0	2 / 2.8	0 / 0.0	0 / 0.0	9 / 5.3
HOMEMAKER	3 / 6.2	3 / 14.3	7 / 9.7	0 / 0.0	1 / 5.9	14 / 8.2
COLUMN TOTAL	48 / 28.2	21 / 12.4	72 / 42.3	12 / 7.1	17 / 10.0	170 / 100.0

23 OUT OF 35 (65.7%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.136
CHI SQUARE = 48.57385 WITH 24 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0021

NUMBER OF MISSING OBSERVATIONS = 19

Appendix G.IX. Age by Control Treatments

C R O S S T A B U L A T I O N O F
AGE- G) AGE NOW
BY CASEVAR- CONTROL STUDY VARIATION TESTED

	CASEVAR					
COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
AGE						
10-17.	6	2	1	11	2	22
	11.1	9.1	1.4	100.0	11.1	12.3
18-28.	17	16	32	0	11	76
	31.5	72.7	43.2	0.0	61.1	42.5
30-49.	16	2	15	0	5	38
	29.6	9.1	20.3	0.0	27.8	21.2
50-83.	15	2	26	0	0	43
	27.8	9.1	35.1	0.0	0.0	24.0
COLUMN TOTAL	54	22	74	11	18	179
	30.2	12.3	41.3	5.1	10.1	100.0

8 OUT OF 20 (40.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 1.308
CHI SQUARE = 116.93066 WITH 12 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 10

Appendix G.X. Education by Control Treatments

C R O S S T A B U L A T I O N O F
EDUCATN- H) AGE WHEN FULLTIME EDUCATION STOPPED
BY CASEVAR- CONTROL STUDY VARIATION TESTED

	CASEVAR					
COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
EDUCATN						
"NEVER",NEVER WAS	1	0	0	0	1	2
	1.9	0.0	0.0	0.0	5.6	1.2
STILL STUDENT	7	7	23	12	5	54
	13.5	38.9	34.8	100.0	27.8	32.5
8-18.	26	3	12	0	5	46
	50.0	16.7	18.2	0.0	27.8	27.7
19-89.	18	8	31	0	7	64
	34.6	44.4	47.0	0.0	38.9	38.6
COLUMN TOTAL	52	18	66	12	18	166
	31.3	10.9	39.8	7.2	10.8	100.0

9 OUT OF 20 (45.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.140
CHI SQUARE = 49.39075 WITH 12 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 23

Appendix G.XI. Prior Knowledge of Testing by Control Treatments

C R O S S T A B U L A T I O N O F
KNEWASK- I) PRIOR KNOWLEDGE OF QUESTIONNAIRE
BY CASEVAR- CONTROL STUDY VARIATION TESTED

	CASEVAR					
COUNT / COL PCT	PILOT UNSEEN	PRE-POST GROUPS	ASK TO STUDY	FIRHILL ASKSTUDY	TEST YOURSELF	ROW TOTAL
KNEWASK						
NO	19	22	69	8	18	136
	33.3	100.0	94.5	72.7	100.0	75.1
YES-SAW\KNEW\HEA	0	0	4	3	0	7
	0.0	0.0	5.5	27.3	0.0	3.9
CONTROL, NEVER SAW	38	0	0	0	0	38
	66.7	0.0	0.0	0.0	0.0	21.0
COLUMN TOTAL	57	22	73	11	18	181
	31.5	12.2	40.3	6.1	9.9	100.0

7 OUT OF 15 (46.7%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.585
CHI SQUARE = 122.04172 WITH 8 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 8

Appendix H. Day of Week- Full Table for Main Display Treatments:

C R O S S T A B U L A T I O N O F
DAYOWEEK- DAY OF THE WEEK TESTING TOOK PLACE
BY CASEVAR- DISPLAY CASE TREATMENT VARIATION TESTED

DAYOWEEK	CASEVAR: IMT-MULTITRACK	MTG-MULTITRACK,	BT-BLOCK LABELS,	BTG-BLOC K LABELS	BU-BLOCK LABELS,	BUG-BLOC K LABELS	ROW TOTAL
SATURDAY	52	23	10	6	25	20	136
							22.7
SUNDAY	14	8	0	13	7	28	70
							11.7
MONDAY	17	0	30	6	10	21	84
							14.0
TUESDAY	0	24	15	28	0	15	82
							13.7
WEDNESDAY	0	11	18	16	21	13	79
							13.2
THURSDAY	5	13	8	17	8	3	54
							9.0
FRIDAY	12	21	19	14	29	0	95
							15.8
COLUMN TOTAL	100	100	100	100	100	100	600
	16.7	16.7	16.7	16.7	16.7	16.7	100.0

CHI SQUARE = 254.30659 WITH 30 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

Appendix I. Comparison of Weekend vs Weekday Visitor Characteristics for the Main Display Treatments:

```
CROSSTABS       =DAYOWEEK BY
                FIVISIT,TIMESYR,INTRST,KNOWLDG,RESIDENC,SEX,OCCSEG
                AGE,EDUCATN,KNEWASK,TCORECT
```

Appendix I.I. Weekend and Weekday by First Visit

```
C R O S S T A B U L A T I O N   O F
    DAYOWEEK-  DAY OF THE WEEK TESTING TOOK PLACE
      BY  FIVISIT-   A1) IS THIS YOUR FIRST VISIT TO THIS MUSEUM?


               FIVISIT
       COUNT   I
      ROW PCT IYES       NO          ROW
              I                     TOTAL
              I         I         I
DAYOWEEK------I---------I---------I
              I   147   I    44   I   191
  WEEKEND     I  77.0   I  23.0   I  35.2
             -I---------I---------I
              I   313   I    38   I   351
  WEEKDAY     I  89.2   I  10.8   I  64.8
             -I---------I---------I
      COLUMN      460        82       542
       TOTAL     84.9       15.1     100.0


CORRECTED CHI SQUARE =    13.42730 WITH 1 DEGREE OF FREEDOM.
SIGNIFICANCE =  0.0002
    RAW CHI SQUARE =    14.36255 WITH 1 DEGREE OF FREEDOM.
    SIGNIFICANCE =  0.0002

NUMBER OF MISSING OBSERVATIONS =      58
```

Appendix I.II. Weekend and Weekday by Number of Visits in Past Year

C R O S S T A B U L A T I O N O F
DAYOWEEK- DAY OF THE WEEK TESTING TOOK PLACE
BY TIMESYR- A2) NO.OF VISITS IN PAST YEAR

	TIMESYR						
COUNT / ROW PCT	1	2	3	4	5	6	ROW TOTAL
WEEKEND	9	6	1	0	1	1	18
	50.0	33.3	5.6	0.0	5.6	5.6	46.2
WEEKDAY	10	7	1	1	0	2	21
	47.6	33.3	4.8	4.8	0.0	9.5	53.8
COLUMN TOTAL	19	13	2	1	1	3	39
	48.7	33.3	5.1	2.6	2.6	7.7	100.0

8 OUT OF 12 (66.7%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.462
CHI SQUARE = 2.24540 WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE = 0.8143

NUMBER OF MISSING OBSERVATIONS = 561

Appendix I.III. Weekend and Weekday by Interest in Scottish History

```
C R O S S T A B U L A T I O N   O F
    DAYOWEEK-  DAY OF THE WEEK TESTING TOOK PLACE
      BY  INTRST-    B)  AVOWED INTEREST IN EARLY SCOTTISH HISTORY

                  INTRST
           COUNT  I
           ROW PCT IVERY INT MODERATE SLIGHT-     ROW
                   IERESTED  LY       NONE       TOTAL
                   I         I        I         I
DAYOWEEK   --------I---------I--------I---------I
                   I    34   I   101  I    55   I   190
  WEEKEND          I  17.9   I  53.2  I  28.9   I  35.1
                  -I---------I--------I---------I
                   I    70   I   169  I   112   I   351
  WEEKDAY          I  19.9   I  48.1  I  31.9   I  64.9
                  -I---------I--------I---------I
           COLUMN      104       270       167       541
            TOTAL     19.2      49.9      30.9     100.0

CHI SQUARE =     1.23918 WITH   2 DEGREES OF FREEDOM
SIGNIFICANCE =  0.5382
NUMBER OF MISSING OBSERVATIONS =      59
```

Appendix I.IV. Weekend and Weekday by Knowledge of Scottish Archaeology

```
C R O S S T A B U L A T I O N   O F
    DAYOWEEK-  DAY OF THE WEEK TESTING TOOK PLACE
      BY  KNOWLDG-   C)  GEN.KNOWLEDGE OF SCOTTISH ARCHAEOLOGY

                  KNOWLDG
           COUNT  I
           ROW PCT IFAIR-V.G POOR-NO    ROW
                   IOOD KNOW KNOWLEDG  TOTAL
                   I         I         I
DAYOWEEK   --------I---------I---------I
                   I    57   I   133   I   190
  WEEKEND          I  30.0   I  70.0   I  35.1
                  -I---------I---------I
                   I    98   I   254   I   352
  WEEKDAY          I  27.8   I  72.2   I  64.9
                  -I---------I---------I
           COLUMN      155       387       542
            TOTAL     28.6      71.4     100.0

CORRECTED CHI SQUARE =      0.18589 WITH 1 DEGREE OF FREEDOM.
SIGNIFICANCE =  0.6664
      RAW CHI SQUARE =      0.28171 WITH 1 DEGREE OF FREEDOM.
      SIGNIFICANCE =  0.5956
NUMBER OF MISSING OBSERVATIONS =      58
```

Appendix I.V. Weekend and Weekday by Residence

```
C R O S S T A B U L A T I O N   O F
    DAYOWEEK-  DAY OF THE WEEK TESTING TOOK PLACE
      BY  RESIDENC-  D)  RESIDENCE

                  RESIDENC
           COUNT  I
           ROW PCT IEDINBURG SCOTLAND ENGLAND  U.S.A    UNSPECIF   ROW
                  IH                                    IED ALIE  TOTAL
                  I        I        I        I        I        I
DAYOWEEK   --------I--------I--------I--------I--------I--------I
                  I    46  I    33  I    22  I    80  I     6  I   187
  WEEKEND         I  24.6  I  17.6  I  11.8  I  42.8  I   3.2  I  34.9
                 -I--------I--------I--------I--------I--------I
                  I    27  I    39  I    47  I   217  I    19  I   349
  WEEKDAY         I   7.7  I  11.2  I  13.5  I  62.2  I   5.4  I  65.1
                 -I--------I--------I--------I--------I--------I
           COLUMN       73       72       69      297       25      536
            TOTAL     13.6     13.4     12.9     55.4      4.7    100.0

CHI SQUARE =     39.06413 WITH    4 DEGREES OF FREEDOM
SIGNIFICANCE =  0.0000

NUMBER OF MISSING OBSERVATIONS =     64
```

Appendix I.VI. Weekend and Weekday by Sex

```
C R O S S T A B U L A T I O N   O F
    DAYOWEEK-  DAY OF THE WEEK TESTING TOOK PLACE
      BY  SEX

                  SEX
           COUNT  I
           ROW PCT IMALE     FEMALE     ROW
                  I                    TOTAL
                  I        I        I
DAYOWEEK   --------I--------I--------I
                  I   106  I    99  I   205
  WEEKEND         I  51.7  I  48.3  I  34.3
                 -I--------I--------I
                  I   202  I   190  I   392
  WEEKDAY         I  51.5  I  48.5  I  65.7
                 -I--------I--------I
           COLUMN      308      289      597
            TOTAL     51.6     48.4    100.0

CORRECTED CHI SQUARE =      0.00000 WITH 1 DEGREE OF FREEDOM.
SIGNIFICANCE =  1.0000
      RAW CHI SQUARE =      0.00168 WITH 1 DEGREE OF FREEDOM.
      SIGNIFICANCE =  0.9673

NUMBER OF MISSING OBSERVATIONS =        3
```

Appendix I.VII. Weekend and Weekday by Occupation

C R O S S T A B U L A T I O N O F
DAYOWEEK- DAY OF THE WEEK TESTING TOOK PLACE
BY OCCSEG- F1) OCCUPATION-SOCIAL ECONOMIC GROUP

	OCCSEG							
COUNT / ROW PCT	PROFESIONAL	MANAGERS	SKILED	PARTLY SKILED	UNSKILLED	STUDNT RET,UN EMP,TR	HOME-MAKER	ROW TOTAL
WEEKEND	36	51	21	4	1	32	16	22	183
	19.7	27.9	11.5	2.2	0.5	17.5	8.7	12.0	35.6
WEEKDAY	60	102	32	3	1	57	37	39	331
	18.1	30.8	9.7	0.9	0.3	17.2	11.2	11.8	64.4
COLUMN TOTAL	96	153	53	7	2	89	53	61	514
	18.7	29.8	10.3	1.4	0.4	17.3	10.3	11.9	100.0

4 OUT OF 16 (25.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.712
CHI SQUARE = 3.15346 WITH 7 DEGREES OF FREEDOM
SIGNIFICANCE = 0.8705

NUMBER OF MISSING OBSERVATIONS = 86

Appendix I.VIII. Weekend and Weekday by Age

C R O S S T A B U L A T I O N O F
DAYOWEEK- DAY OF THE WEEK TESTING TOOK PLACE
BY AGE- G) AGE NOW

	AGE				
COUNT / ROW PCT	10-17	18-29	30-49	50-83	ROW TOTAL
DAYOWEEK					
WEEKEND	3	64	66	52	185
	1.6	34.6	35.7	28.1	34.9
WEEKDAY	11	114	97	123	345
	3.2	33.0	28.1	35.7	65.1
COLUMN TOTAL	14	178	163	175	530
	2.6	33.6	30.8	33.0	100.0

1 OUT OF 8 (12.5%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 4.887
CHI SQUARE = 5.51887 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.1375

NUMBER OF MISSING OBSERVATIONS = 70

Appendix I.IX. Weekend and Weekday by Education

```
C R O S S T A B U L A T I O N   O F
    DAYOWEEK-  DAY OF THE WEEK TESTING TOOK PLACE
      BY  EDUCATN-   H)  AGE WHEN FULLTIME EDUCATION STOPPED

                  EDUCATN
          COUNT  I
          ROW PCT I"NEVER", STILL    BEFORE19 AFTER18     ROW
                  INEVER WA STUDENT                      TOTAL
                  I        I        I   8-18.I  19-89.I
DAYOWEEK  --------I--------I--------I--------I--------I
                  I     2  I    33  I    32  I   101  I   168
  WEEKEND         I   1.2  I  19.6  I  19.0  I  60.1  I  34.1
                 -I--------I--------I--------I--------I
                  I     5  I    57  I    70  I   193  I   325
  WEEKDAY         I   1.5  I  17.5  I  21.5  I  59.4  I  65.9
                 -I--------I--------I--------I--------I
          COLUMN        7       90      102      294      493
           TOTAL      1.4     18.3     20.7     59.6    100.0

    2 OUT OF     8 ( 25.0%) OF THE VALID CELLS HAVE
    EXPECTED CELL FREQUENCY LESS THAN 5.0.
       MINIMUM EXPECTED CELL FREQUENCY =  2.385
CHI SQUARE =     0.70524 WITH   3 DEGREES OF FREEDOM
SIGNIFICANCE =  0.8720

NUMBER OF MISSING OBSERVATIONS =     107
```

Appendix I.X. Weekend and Weekday by Prior Knowledge of Testing

```
C R O S S T A B U L A T I O N   O F
    DAYOWEEK- DAY OF THE WEEK TESTING TOOK PLACE
      BY  KNEWASK-   I)  PRIOR KNOWLEDGE OF QUESTIONNAIRE

                  KNEWASK
           COUNT  I
          ROW PCT INO       YES,AWAR   ROW
                  I         E OF TES  TOTAL
                  I        I        I
DAYOWEEK   --------I--------I--------I
                  I   186  I     1  I   187
  WEEKEND         I  99.5  I   0.5  I  35.1
                 -I--------I--------I
                  I   345  I     1  I   346
  WEEKDAY         I  99.7  I   0.3  I  64.9
                 -I--------I--------I
          COLUMN      531        2      533
           TOTAL     99.6      0.4    100.0

    2 OUT OF     4 ( 50.0%) OF THE VALID CELLS HAVE
    EXPECTED CELL FREQUENCY LESS THAN 5.0.
       MINIMUM EXPECTED CELL FREQUENCY =  0.702
CORRECTED CHI SQUARE =     0.00000 WITH 1 DEGREE OF FREEDOM.
SIGNIFICANCE =  1.0000
      RAW CHI SQUARE =     0.19610 WITH 1 DEGREE OF FREEDOM.
      SIGNIFICANCE =  0.6579

NUMBER OF MISSING OBSERVATIONS =      67
```

Appendix I.XI. Weekend and Weekday by Test Score

C R O S S T A B U L A T I O N O F
DAYOWEEK- DAY OF THE WEEK TESTING TOOK PLACE
BY TCORECT- CORRECT TEST SCORE

		TCORECT					
	COUNT / ROW PCT	0.	1-3.	4-6.	7-9.	10-18.	ROW TOTAL
DAYOWEEK							
WEEKEND		19	47	80	45	15	206
		9.2	22.8	38.8	21.8	7.3	34.3
WEEKDAY		46	103	142	68	35	394
		11.7	26.1	36.0	17.3	8.9	65.7
	COLUMN TOTAL	65	150	222	113	50	600
		10.8	25.0	37.0	18.8	8.3	100.0

CHI SQUARE = 3.56180 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.4685

Appendix J. Proportion of First Time Visitors in M.D.T. Study, Control Groups:

NOTE: M.D.T.= MAIN DISPLAY TREATMENTS

C R O S S T A B U L A T I O N O F
CASEVAR- MAIN DISPLAY CASE TREATMENT VARIATION TESTED
BY FIVISIT- A1) IS THIS YOUR FIRST VISIT TO THIS MUSEUM

	FIVISIT		
COUNT / ROW PCT	YES	NO	ROW TOTAL
CASEVAR			
MT	69 / 75.8	22 / 24.2	91 / 16.8
MTG	83 / 87.4	12 / 12.6	95 / 17.5
BT	79 / 90.8	8 / 9.2	87 / 16.1
BTG	82 / 90.1	9 / 9.9	91 / 16.8
BU	78 / 89.7	9 / 10.3	87 / 16.1
BUG	69 / 75.8	22 / 24.2	91 / 16.8
COLUMN TOTAL	460 / 84.9	82 / 15.1	542 / 100.0

CHI SQUARE = 17.94377 WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0030

NUMBER OF MISSING OBSERVATIONS = 58

C R O S S T A B U L A T I O N O F
CASEVAR- CONTROL STUDY VARIATION TESTED
BY FIVISIT A1) IS THIS YOUR FIRST VISIT TO THIS MUSEUM?

CASEVAR	COUNT / ROW PCT	FIVISIT YES	FIVISIT NO	ROW TOTAL
PILOT UNSEEN		34	22	56
		60.7	39.3	30.8
PRE-POST GROUPS		21	1	22
		95.5	4.5	12.1
ASK TO STUDY		61	13	74
		82.4	17.6	40.6
FIRHILL ASKSTUDY		10	2	12
		83.3	16.7	6.6
TEST YOURSELF		16	2	18
		88.9	11.1	9.9
COLUMN TOTAL		142	40	182
		78.0	22.0	100.0

2 OUT OF 10 (20.0%) OF THE VALID CELLS HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 2.513
CHI SQUARE = 18.95007 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0008

NUMBER OF MISSING OBSERVATIONS = 7

Appendix K. Sex Ratios in M.D.T. Sample, Control Groups, and All Gallery Visitors:

	ALL GALLERY VISITORS:				TESTED SAMPLE:	
CASEVAR	AllMales	AllFems.	AllChild	AllVis.	MaleSubj	FemSubj
MT	200	161	73	434	48	51
MTG	229	199	77	505	53	46
BT	207	206	34	447	47	53
BTG	245	233	22	500	50	49
BU	236	191	29	456	51	49
BUG	173	164	34	371	59	41
TOTAL	1290	1154	269	2713	308	289
%	(47.6%)	(42.5%)	(9.9%)	(100%)	(51.6%)	(48.4%)

CONTROL GROUPS:

Pilot					30(57.7%)	22(42.3%)
Pre-Post					9(40.9%)	13(59.1%)
Ask to Study					36(48%)	39(52%)
Firhill (ask)					12(100%)	0
Test Yourself					9(50%)	9(50%)
Column Total:				Total=179	96(53.6%)	88(46.4%)
MTa	309	338	127	774	52	46
RMT	291	268	82	641	50	49

Appendix L. Knowledge Assessments of M.D.T. Sample Compared with Control Groups:

C R O S S T A B U L A T I O N O F
CASEVAR- MAIN DISPLAY CASE TREATMENT VARIATION TESTED
BY KNOWLDG- C) GEN.KNOWLEDGE OF SCOTTISH ARCHAEOLOGY

	KNOWLDG		
COUNT / ROW PCT	FAIR-V.G KNOWLEDG	POOR-NO KNOWLEDG	ROW TOTAL
CASEVAR			
MT	31 / 33.7	61 / 66.3	92 / 17.0
MTG	25 / 26.6	69 / 73.4	94 / 17.3
BT	30 / 34.9	56 / 65.1	86 / 15.9
BTG	24 / 26.1	68 / 73.9	92 / 17.0
BU	20 / 23.0	67 / 77.0	87 / 16.1
BUG	25 / 27.5	66 / 72.5	91 / 16.8
COLUMN TOTAL	155 / 28.6	387 / 71.4	542 / 100.0

CHI SQUARE = 4.70061 WITH 5 DEGREES OF FREEDOM
SIGNIFICANCE = 0.4535

NUMBER OF MISSING OBSERVATIONS = 58

C R O S S T A B U L A T I O N O F
CASEVAR- CONTROL STUDY VARIATION TESTED
BY KNOWLDG- C) GEN.KNOWLEDGE OF SCOTTISH ARCHAEOLOGY

	KNOWLDG		
COUNT / ROW PCT	FAIR-V.G KNOWLEDG	POOR-NO KNOWLEDG	ROW TOTAL
CASEVAR			
PILOT UNSEEN	14 / 26.4	39 / 73.6	53 / 29.4
PRE-POST GROUPS	5 / 22.7	17 / 77.3	22 / 12.2
ASK TO STUDY	31 / 41.3	44 / 58.7	75 / 41.7
FIRHILL ASKSTUDY	8 / 66.7	4 / 33.3	12 / 6.7
TEST YOURSELF	11 / 61.1	7 / 38.9	18 / 10.0
COLUMN TOTAL	69 / 38.3	111 / 61.7	180 / 100.0

1 OUT OF 10 (10.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 4.508
CHI SQUARE = 14.81918 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0051

NUMBER OF MISSING OBSERVATIONS = 9

Appendix M. Age Bands- Comparison of M.D.T. Sample with Control Groups:

C R O S S T A B U L A T I O N O F
CASEVAR- MAIN DISPLAY CASE TREATMENT VARIATION TESTED
BY AGE- G) AGE NOW

		AGE				
	COUNT / ROW PCT	10-17.	18-29.	30-49.	50-83.	ROW TOTAL
CASEVAR						
MT		3	29	26	30	88
		3.4	33.0	29.5	34.1	16.6
MTG		4	30	35	25	94
		4.3	31.9	37.2	26.6	17.7
BT		2	29	16	39	86
		2.3	33.7	18.6	45.3	16.2
BTG		2	32	27	26	87
		2.3	36.8	31.0	29.9	16.4
BU		3	27	33	23	86
		3.5	31.4	38.4	26.7	16.2
BUG		0	31	26	32	89
		0.0	34.8	29.2	36.0	16.8
	COLUMN TOTAL	14	178	163	175	530
		2.6	33.6	30.8	33.0	100.0

6 OUT OF 24 (25.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 2.272
CHI SQUARE = 18.12419 WITH 15 DEGREES OF FREEDOM
SIGNIFICANCE = 0.2562

NUMBER OF MISSING OBSERVATIONS = 70

C R O S S T A B U L A T I O N O F
CASEVAR- CONTROL STUDY VARIATION TESTED
BY AGE- G) AGE NOW

CASEVAR	COUNT / ROW PCT	AGE 10-17.	AGE 18-28.	AGE 30-49.	AGE 50-83.	ROW TOTAL
PILOT UNSEEN		6	17	16	15	54
		11.1	31.5	29.6	27.8	30.2
PRE-POST GROUPS		2	16	2	2	22
		9.1	72.7	9.1	9.1	12.3
ASK TO STUDY		1	32	15	26	74
		1.4	43.2	20.3	35.1	41.3
FIRHILL ASKSTUDY		11	0	0	0	11
		100.0	0.0	0.0	0.0	5.1
TEST YOURSELF		2	11	5	0	18
		11.1	61.1	27.8	0.0	10.1
	COLUMN TOTAL	22	76	38	43	179
		12.3	42.5	21.5	24.0	100.0

8 OUT OF 20 (40.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 1.308
CHI SQUARE = 116.93069 WITH 12 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 10

Appendix N. Comparison of Residence with Interest and Knowledge Assessments **(M.D.T.)**

RESIDENC	INTRST	VERY INTERESTED	MODERATELY	SLIGHTLY -NOT INT	ROW TOTAL
SCOTLAND	COUNT	39	75	30	144
	ROW PCT	27.1	52.1	20.8	27.1
	COL PCT	38.2	28.1	18.4	
ENGLAND	COUNT	12	35	22	69
	ROW PCT	17.4	50.7	31.9	13.0
	COL PCT	11.8	13.1	13.5	
U.S.A	COUNT	48	145	101	294
	ROW PCT	16.3	49.3	34.4	55.3
	COL PCT	47.1	54.3	62.0	
UNSPECIFIED ALIEN	COUNT	3	12	10	25
	ROW PCT	12.0	48.0	40.0	4.7
	COL PCT	2.9	4.5	6.1	
COLUMN TOTAL		102 19.2	267 50.2	163 30.6	532 100.0

1 OUT OF 12 (8.3%) OF THE VALID CELLS HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 4.793
CHI SQUARE = 13.49526 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0358

NUMBER OF MISSING OBSERVATIONS = 68

RESIDENC	KNOWLDG: NONE-POOR	FAIR-V.GOOD KNOW	ROW TOTAL
SCOTLAND	78	64	142
	54.9	45.1	26.7
	20.6	41.8	
ENGLAND	36	33	69
	52.2	47.8	13.0
	9.5	21.6	
U.S.A	240	56	296
	81.1	18.9	55.6
	63.3	36.6	
UNSPECIFIED ALIEN	25	0	25
	100.0	0.0	4.7
	6.6	0.0	
COLUMN TOTAL	379	153	532
	71.2	28.8	100.0

CHI SQUARE = 54.76465 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 68

Appendix O. Comparison of Test Score with Interest in, and Knowledge of, Archaeology (M.D.T.):

	INTRST			
COUNT / ROW PCT / COL PCT	VERY INTERESTED	MODERATELY	SLIGHTLY -NOT INT	ROW TOTAL
TCORECT				
0.	2	6	14	22
	9.1	27.3	63.6	4.1
	1.9	2.2	8.4	
1-3.	26	64	48	138
	18.8	46.4	34.8	25.5
	25.0	23.7	28.7	
4-6.	36	111	73	220
	16.4	50.5	33.2	40.7
	34.6	41.1	43.7	
7-9.	22	64	25	111
	19.8	57.7	22.5	20.5
	21.2	23.7	15.0	
10-18.	18	25	7	50
	36.0	50.0	14.0	9.2
	17.3	9.3	4.2	
COLUMN TOTAL	104	270	167	541
	19.2	49.9	30.9	100.0

1 OUT OF 15 (6.7%) OF THE VALID CELLS HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 4.229
CHI SQUARE = 29.24715 WITH 8 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0003

NUMBER OF MISSING OBSERVATIONS = 59

TCORECT	KNOWLDG		
COUNT / ROW PCT / COL PCT	NONE-POOR	FAIR-V.GOOD	ROW TOTAL
0.	21	2	23
	91.3	8.7	4.2
	5.4	1.3	
1-3.	103	36	139
	74.1	25.9	25.6
	26.6	23.2	
4-6.	159	61	220
	72.3	27.7	40.6
	41.1	39.4	
7-9.	74	36	110
	67.3	32.7	20.3
	19.1	23.2	
10-18.	30	20	50
	60.0	40.0	9.2
	7.8	12.9	
COLUMN TOTAL	387	155	542
	71.4	28.6	100.0

CHI SQUARE = 9.14099 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0577

NUMBER OF MISSING OBSERVATIONS = 58

C R O S S T A B U L A T I O N O F
TCORECT- TOTAL CORRECT TEST SCORE
BY INTRST- B) INTEREST IN EARLY SCOTTISH HISTORY
CONTROLLING FOR..
KNOWLDG- C) GEN.KNOWLEDGE OF SCOTTISH ARCHAEOLOGY

VALUE = FAIR to VERY GOOD KNOWLEDGE:

TCORECT	COUNT / ROW PCT	INTRST VERY INTERESTED	MODERATELY	SLIGHT-NONE	ROW TOTAL
0.		1	1	0	2
		50.0	50.0	0.0	1.3
1-3.		11	16	9	36
		30.6	44.4	25.0	23.2
4-6.		19	34	8	61
		31.1	55.7	13.1	39.4
7-9.		13	20	3	36
		36.1	55.6	8.3	23.2
10-18.		11	9	0	20
		55.0	45.0	0.0	12.9
COLUMN TOTAL		55	80	20	155
		35.5	51.6	12.9	100.0

6 OUT OF 15 (40.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.258
CHI SQUARE = 11.18374 WITH 8 DEGREES OF FREEDOM
SIGNIFICANCE = 0.1915

C R O S S T A B U L A T I O N O F
TCORECT- TOTAL CORRECT TEST SCORES
BY INTRST- B) INTEREST IN EARLY SCOTTISH HISTORY
CONTROLLING FOR..
KNOWLDG- C) GENERAL KNOWLEDGE OF SCOTTISH ARCHAEOLOGY

VALUE = POOR to NO KNOWLEDGE of Scottish Archaeology:

TCORECT	COUNT / ROW PCT	INTRST: VERY INTERESTED	MODERATELY	SLIGHT-NONE	ROW TOTAL
0.		1	5	14	20
		5.0	25.0	70.0	5.2
1-3.		15	48	39	102
		14.7	47.1	38.2	26.6
4-6.		17	76	65	158
		10.8	48.1	41.1	41.3
7-9.		7	44	22	73
		9.6	60.3	30.1	19.1
10-18.		7	16	7	30
		23.3	53.3	23.3	7.8
COLUMN TOTAL		47	189	147	383
		12.3	49.3	38.4	100.0

2 OUT OF 15 (13.3%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 2.454
CHI SQUARE = 18.07729 WITH 8 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0207

NUMBER OF MISSING OBSERVATIONS = 62

Appendix P. Test Score Compared with Selected Visitor Characteristics, for the M.D.T.

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY INTRST- B) INTEREST IN EARLY SCOTTISH HISTORY

	INTRST		
COUNT / COL PCT	MODERATELY	SLIGHT-NONE	ROW TOTAL
TCORECT			
0-3.	98	62	160
	26.2	37.1	29.6
4-6.	147	73	220
	39.3	43.7	40.7
7-9.	86	25	111
	23.0	15.0	20.5
10-18.	43	7	50
	11.5	4.2	9.2
COLUMN TOTAL	374	167	541
	69.1	30.9	100.0

CHI SQUARE = 15.49921 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0014

NUMBER OF MISSING OBSERVATIONS = 59

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY KNOWLDG C) GEN.KNOWLEDGE OF SCOTTISH ARCHAEOLOG

	KNOWLDG		
COUNT / COL PCT	POOR-NONE	FAIR-V.GOOD KNOW	ROW TOTAL
TCORECT			
0-3.	124	38	162
	32.0	24.5	29.9
4-6.	159	61	220
	41.1	39.4	40.6
7-9.	74	36	110
	19.1	23.2	20.3
10-18.	30	20	50
	7.8	12.9	9.2
COLUMN TOTAL	387	155	542
	71.4	28.6	100.0

CHI SQUARE = 6.28060 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0987

NUMBER OF MISSING OBSERVATIONS = 58

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY INTERACT- TYPE OF VISITOR INTERACTION

	INTERACT				
COUNT / COL PCT	IMUCH TALK	SPECIFIC AGE	LITTLE TALK	NOTES OR GESTURES	ROW TOTAL
TCORECT					
0-3.	22	12	8	4	46
	29.3	41.4	26.7	57.1	32.6
4-6.	30	10	10	1	51
	40.0	34.5	33.3	14.3	36.2
7-9.	16	5	6	1	28
	21.3	17.2	20.0	14.3	19.9
10-18.	7	2	6	1	16
	9.3	6.9	20.0	14.3	11.3
COLUMN TOTAL	75	29	30	7	141
	53.2	20.6	21.3	5.0	100.0

6 OUT OF 16 (37.5%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.794
CHI SQUARE = 6.96607 WITH 9 DEGREES OF FREEDOM
SIGNIFICANCE = 0.6407

NUMBER OF MISSING OBSERVATIONS = 459

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY DIRDOOR- INITIAL TRAFFIC PATTERN OF SUBJECT

TCORECT	DIRDOOR UNKNOWN	TO CASE	RT,LF OR STRAIGHT	ROW TOTAL
0-3.	56	144	15	215
	38.4	36.1	27.3	35.8
4-6.	59	144	19	222
	40.4	36.1	34.5	37.0
7-9.	22	76	15	113
	15.1	19.0	27.3	18.8
10-18.	9	35	6	50
	6.2	8.8	10.9	8.3
COLUMN TOTAL	146	399	55	600
	24.3	66.5	9.2	100.0

1 OUT OF 12 (8.3%) OF THE VALID CELLS HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 4.583
CHI SQUARE = 6.57208 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.3622

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY DIRCASE- DIRECTION FROM WHICH SUBJECTS APPROACHED DISPLAY CASE

	DIRCASE		
COUNT / COL PCT	RIGHT, TOTAL	LEFT, TOTAL	ROW TOTAL
TCORECT			
0-3.	182	33	215
	36.0	35.1	35.8
4-6.	188	34	222
	37.2	36.2	37.0
7-9.	94	19	113
	18.6	20.2	18.8
10-18.	42	8	50
	8.3	8.5	8.3
COLUMN TOTAL	506	94	600
	84.3	15.7	100.0

CHI SQUARE = 0.15400 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.9846

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY VIEWSEQ- SEQUENCE OF VIEWING AGES IN CASE

	VIEWSEQ			
COUNT / COL PCT	IRT TO LF I,PLUS	LF TO RT ,PLUS	RANDOM	ROW TOTAL
TCORECT				
0-3.	164	30	21	215
	33.5	46.2	45.7	35.8
4-6.	185	20	17	222
	37.8	30.8	37.0	37.0
7-9.	98	8	7	113
	20.0	12.3	15.2	18.8
10-18.	42	7	1	50
	8.6	10.8	2.2	8.3
COLUMN TOTAL	489	65	46	600
	81.5	10.8	7.7	100.0

1 OUT OF 12 (8.3%) OF THE VALID CELLS HAVE EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 3.833
CHI SQUARE = 9.42546 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.1510

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY ENTRYGP- ENTRY GROUP OF SUBJECT

	ENTRYGP				
COUNT / COL PCT	ALONE	2 PEOPLE	3 PEOPLE	4OR MORE	ROW TOTAL
TCORECT					
0-3.	78	100	29	8	215
	35.1	35.8	44.6	23.5	35.8
4-6.	77	110	21	14	222
	34.7	39.4	32.3	41.2	37.0
7-9.	41	50	10	12	113
	18.5	17.9	15.4	35.3	18.8
10-18.	26	19	5	0	50
	11.7	6.8	7.7	0.0	8.3
COLUMN TOTAL	222	279	65	34	600
	37.0	46.5	10.8	5.7	100.0

1 OUT OF 16 (6.2%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 2.833
CHI SQUARE = 16.30220 WITH 9 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0608

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY- TIMODAY TIME OF DAY WHEN CASE VIEWED

TCORECT	COUNT / COL PCT	TIMODAY: MORNING	LUNCH-TIME	EARLY AFTERNOON	ROW TOTAL
0-3.		27	68	120	215
		26.0	38.0	37.9	35.8
4-6.		49	64	109	222
		47.1	35.8	34.4	37.0
7-9.		18	33	62	113
		17.3	18.4	19.6	18.8
10-18.		10	14	26	50
		9.6	7.8	8.2	8.3
COLUMN TOTAL		104	179	317	600
		17.3	29.8	52.8	100.0

CHI SQUARE = 7.45913 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.2805

C R O S S T A B U L A T I O N O F
TCORECT-TEST SCORE
BY VIEWTIME- SECONDS SPENT VIEWING THE DISPLAY CASE

	VIEWTIME							
COUNT I COL PCTI	45"- 129"	130"- 229"	230"- 329"	330"- 499"	500"- 699"	700"- 999"	1000"- 1643"	ROW TOTAL
TCORECT I VIEWED I								
0-3. I	105	77	19	6	5	2	1	215
	46.1	37.7	26.0	10.3	20.8	25.0	20.0	35.8
4-6. I	84	82	23	25	6	2	0	222
	36.8	40.2	31.5	43.1	25.0	25.0	0.0	37.0
7-9. I	31	39	17	15	7	2	2	113
	13.6	19.1	23.3	25.9	29.2	25.0	40.0	18.8
10-18. I	8	6	14	12	6	2	2	50
	3.5	2.9	19.2	20.7	25.0	25.0	40.0	8.3
COLUMN TOTAL	228 38.0	204 34.0	73 12.2	58 9.7	24 4.0	8 1.3	5 0.8	600 100.0

11 OUT OF 28 (39.3%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.417
CHI SQUARE = 85.71481 WITH 18 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0000

C R O S S T A B U L A T I O N O F
TCORECT-TEST SCORE
BY GOALTIME SECONDS SPENT READING SIGN

	GOALTIME				
COUNT / COL PCT	5 TO 10 SECONDS	12 TO 20 SECONDS	21 TO 70 SECONDS	READ AFTER VIEW	ROW TOTAL
TCORECT					
0-3.	27	21	8	1	57
	39.1	29.2	30.8	50.0	33.7
4-6.	24	33	9	0	66
	34.8	45.8	34.6	0.0	39.1
7-9.	10	16	6	1	33
	14.5	22.2	23.1	50.0	19.5
10-18.	8	2	3	0	13
	11.6	2.8	11.5	0.0	7.7
COLUMN TOTAL	69	72	26	2	169
	40.8	42.6	15.4	1.2	100.0

5 OUT OF 16 (31.2%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 0.154
CHI SQUARE = 9.91017 WITH 9 DEGREES OF FREEDOM
SIGNIFICANCE = 0.3578
NUMBER OF MISSING OBSERVATIONS = 431

C R O S S T A B U L A T I O N O F
TCORECT-TEST SCORE
BY DAYOWEEK- DAY OF THE WEEK TESTING TOOK PLACE

	DAYOWEEK		
COUNT / COL PCT	WEEKEND	WEEKDAY	ROW TOTAL
TCORECT			
0-3.	66	149	215
	32.0	37.8	35.8
4-6.	80	142	222
	38.8	36.0	37.0
7-9.	45	68	113
	21.8	17.3	18.8
10-18.	15	35	50
	7.3	8.9	8.3
COLUMN TOTAL	206	394	600
	34.3	65.7	100.0

CHI SQUARE = 3.47288 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.3243

C R O S S T A B U L A T I O N O F
TCORECT-TEST SCORE
BY FLORCOND- FLOOR CONDITIONS AT THE START OF THE DAILY EXPERIMENT

TCORECT	FLORCOND COUNT I COL PCT I QUIET	NOISY	ROW TOTAL
0-3.	154	61	215
	35.5	36.7	35.8
4-6.	165	57	222
	38.0	34.3	37.0
7-9.	82	31	113
	18.9	18.7	18.8
10-18.	33	17	50
	7.6	10.2	8.3
COLUMN TOTAL	434 72.3	166 27.7	600 100.0

CHI SQUARE = 1.49843 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.6826

C R O S S T A B U L A T I O N O F
TCORECT-TEST SCORE
BY WEATHER- WEATHER CONDITIONS AT THE START OF THE DAILY EXPERIMENT

	WEATHER			
COUNT / COL PCT	CLOUDY	RAIN	SUN	ROW TOTAL
TCORECT				
0-3.	72 / 35.5	23 / 35.4	120 / 36.1	215 / 35.8
4-6.	77 / 37.9	20 / 30.8	125 / 37.7	222 / 37.0
7-9.	43 / 21.2	12 / 18.5	58 / 17.5	113 / 18.8
10-18.	11 / 5.4	10 / 15.4	29 / 8.7	50 / 8.3
COLUMN TOTAL	203 / 33.8	65 / 10.8	332 / 55.3	600 / 100.0

CHI SQUARE = 7.72677 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.2588

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY CASEVAR- PRINCIPAL DISPLAY TREATMENT TESTED

	CASEVAR			
COUNT / COL PCT	IMTs	BTs	BUs	ROW TOTAL
TCORECT				
0-3.	58	89	68	215
	29.0	44.5	34.0	35.8
4-6.	78	71	73	222
	39.0	35.5	36.5	37.0
7-9.	46	25	42	113
	23.0	12.5	21.0	18.8
10-18.	18	15	17	50
	9.0	7.5	8.5	8.3
COLUMN TOTAL	200	200	200	600
	33.3	33.3	33.3	100.0

CHI SQUARE = 14.21916 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0273

C R O S S T A B U L A T I O N O F
TCORECT- TEST SCORE
BY CASEVAR- WITH GOAL OR WITHOUT GOAL DISPLAY CASE TREATMENT

	CASEVAR		
COUNT / COL PCT	WITHOUT GOALS	WITH GOALS	ROW TOTAL
TCORECT			
0-3.	112 / 37.3	103 / 34.3	215 / 35.8
4-6.	110 / 36.7	112 / 37.3	222 / 37.0
7-9.	56 / 18.7	57 / 19.0	113 / 18.8
10-18.	22 / 7.3	28 / 9.3	50 / 8.3
COLUMN TOTAL	300 / 50.0	300 / 50.0	600 / 100.0

CHI SQUARE = 1.12361 WITH 3 DEGREES OF FREEDOM
SIGNIFICANCE = 0.7714

Appendix Q. Written Comment Score Compared with the Principal Display Treatments and Selected Visitor Characteristics, for the M.D.T.

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY CASEVAR- PRINCIPAL DISPLAY TREATMENT VARIATION TESTED

SUBPOS	COUNT / COL PCT	CASEVAR IMTs	BTs	BUs	ROW TOTAL
-4 TO -1		53	77	72	202
		28.3	41.6	37.3	35.8
1 TO 3		91	76	87	254
		48.7	41.1	45.1	45.0
4 TO 11		43	32	34	109
		23.0	17.3	17.6	19.3
	COLUMN TOTAL	187	185	193	565
		33.1	32.7	34.2	100.0

CHI SQUARE = 7.96623 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0928

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY INTRST- B) INTEREST IN EARLY SCOTTISH HISTORY

SUBPOS	COUNT / COL PCT	INTRST: MODERATELY	INTRST: SLIGHTONE	ROW TOTAL
-4 TO -1		90	61	151
		1	40.7	29.7
1 TO 3		181	68	249
		4	45.3	48.9
4 TO 11		88	21	109
		24.5	14.0	21.4
	COLUMN TOTAL	359	150	509
		70.5	29.5	100.0

CHI SQUARE = 14.69431 WITH 2 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0006

NUMBER OF MISSING OBSERVATIONS = 91

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY KNOWLDG C) GEN.KNOWLEDGE OF SCOTTISH ARCHAEOLOG

	KNOWLDG		
COUNT / COL PCT	POOR-NONE	FAIR-V.GOOD KNOW	ROW TOTAL
SUBPOS	57.I	81.I	
-4 TO -1	114 / 31.4	38 / 25.9	152 / 29.8
1 TO 3	185 / 51.0	64 / 43.5	249 / 48.8
4 TO 11	64 / 17.6	45 / 30.6	109 / 21.4
COLUMN TOTAL	363 / 71.2	147 / 28.8	510 / 100.0

CHI SQUARE = 10.51491 WITH 2 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0052

NUMBER OF MISSING OBSERVATIONS = 90

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY INTERACT TYPE OF VISITOR INTERACTIONS

		INTERACT				
	COUNT					
	COL PCT	IMUCH TAL IK	SPECIFIC AGE	LITTLE T ALK	NOTES OR GESTURE	ROW TOTAL
SUBPOS						
		14	11	12	2	39
-4 TO -1		0	40.7	41.4	28.6	29.3
		37	12	13	4	66
1 TO 3		9	44.4	44.8	57.1	49.6
		19	4	4	1	28
4 TO 11		27.1	14.8	13.8	14.3	21.1
	COLUMN	70	27	29	7	133
	TOTAL	52.6	20.3	21.8	5.3	100.0

3 OUT OF 12 (25.0%) OF THE VALID CELLS HAVE
EXPECTED CELL FREQUENCY LESS THAN 5.0.
MINIMUM EXPECTED CELL FREQUENCY = 1.474
CHI SQUARE = 7.83223 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.2507

NUMBER OF MISSING OBSERVATIONS = 467

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY DIRDOOR INITIAL TRAFFIC PATTERN OF SUBJECT

SUBPOS	COUNT / COL PCT	DIRDOOR UNKNOWN	TO CASE	RT,LF OR STRAIGHT	ROW TOTAL
-4 TO -1		51	134	17	202
		2	35.5	33.3	35.8
1 TO 3		64	165	25	254
		7	43.8	49.0	45.0
4 TO 11		22	78	9	109
		16.1	20.7	17.6	19.3
	COLUMN TOTAL	137	377	51	565
		24.2	66.7	9.0	100.0

CHI SQUARE = 1.76740 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.7784

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY DIRCASE

SUBPOS	DIRCASE: RIGHT,TOTAL	DIRCASE: LEFT,TOTAL	ROW TOTAL
-4 TO -1	170	32	202
	7	36.0	35.8
1 TO 3	209	45	254
	9	50.6	45.0
4 TO 11	97	12	109
	20.4	13.5	19.3
COLUMN TOTAL	476	89	565
	84.2	15.8	100.0

CHI SQUARE = 2.58747 WITH 2 DEGREES OF FREEDOM
SIGNIFICANCE = 0.2742

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY VIEWSEQ SEQUENCE OF VIEWING AGES IN CASE

SUBPOS	COUNT / COL PCT	VIEWSEQ: IRT TO LF, PLUS	LF TO RT, PLUS	RANDOM	ROW TOTAL
-4 TO -1		160	19	23	202
		9	31.1	50.0	35.8
1 TO 3		205	30	19	254
		8	49.2	41.3	45.0
4 TO 11		93	12	4	109
		20.3	19.7	8.7	19.3
	COLUMN TOTAL	458	61	46	565
		81.1	10.8	8.1	100.0

CHI SQUARE = 6.36756 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.1733

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY ENTRYGP ENTRY GROUP OF SUBJECT

SUBPOS	COUNT / COL PCT	ENTRYGP ALONE	2 PEOPLE	3 PEOPLE	4OR MORE	ROW TOTAL
-4 TO -1		75	90	27	10	202
		2	34.0	45.0	30.3	35.8
1 TO 3		82	132	23	17	254
		6	49.8	38.3	51.5	45.0
4 TO 11		50	43	10	6	109
		24.2	16.2	16.7	18.2	19.3
	COLUMN	207	265	60	33	565
	TOTAL	36.6	46.9	10.6	5.8	100.0

CHI SQUARE = 9.62853 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.1412

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY TIMODAY TIME OF DAY WHEN CASE VIEWED

SUBPOS	COUNT / COL PCT	TIMODAY: MORNING	LUNCHTIME	EARLY AF	ROW TOTAL
-4 TO -1		30	69	103	202
		31.6	40.6	34.3	35.8
1 TO 3		37	74	143	254
		38.9	43.5	47.7	45.0
4 TO 11		28	27	54	109
		29.5	15.9	18.0	19.3
	COLUMN	95	170	300	565
	TOTAL	16.8	30.1	53.1	100.0

CHI SQUARE = 9.46289 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0505

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY VIEWTIME TIME SPENT VIEWING CASE

SUBPOS	COUNT / COL PCT	VIEWTIME 45"-129"	130"-229"	230"-329"	330"-499"	500"-699"	700"-999"	1000"-1643"	ROW TOTAL
-4 TO -1		98	63	21	12	7	1	0	202
		4	32.5	30.4	22.2	29.2	12.5	0.0	35.8
1 TO 3		86	94	29	29	8	4	4	254
		8	48.5	42.0	53.7	33.3	50.0	80.0	45.0
4 TO 11		27	37	19	13	9	3	1	109
		12.8	19.1	27.5	24.1	37.5	37.5	20.0	19.3
COLUMN TOTAL		211	194	69	54	24	8	5	565
		37.3	34.3	12.2	9.6	4.2	1.4	0.9	100.0

7 OUT OF 21 (33.3%) OF THE VALID CELLS HAVE EXPECTED CELL FREQUENCY LESS
MINIMUM EXPECTED CELL FREQUENCY = 0.965
CHI SQUARE = 31.65981 WITH 12 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0016

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T IO N O F
SUBPOS- WRITTEN COMMENT SCORE
BY GOALTIME SECONDS SPENT READING SIGN

		GOALTIME				
	COUNT					
	COL PCT	5 TO 10 SECONDS	12 TO 20 SECONDS	21 TO 70 SECONDS	READ AFTER VIEW	ROW TOTAL
SUBPOS						
-4 TO -1		29	27	5	1	62
		6	40.9	20.8	50.0	38.7
1 TO 3		29	30	18	1	78
		6	45.5	75.0	50.0	48.7
4 TO 11		10	9	1	0	20
		14.7	13.6	4.2	0.0	12.5
	COLUMN TOTAL	68	66	24	2	160
		42.5	41.2	15.0	1.2	100.0

4 OUT OF 12 (33.3%) OF THE VALID CELLS HAVE EXPECTED CELL FREQUENCY LESS
MINIMUM EXPECTED CELL FREQUENCY = 0.250
CHI SQUARE = 8.37512 WITH 6 DEGREES OF FREEDOM
SIGNIFICANCE = 0.2119

NUMBER OF MISSING OBSERVATIONS = 440

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY DAYOWEEK DAY OF THE WEEK TESTING TOOK PLACE

SUBPOS	COUNT / COL PCT	DAYOWEEK: WEEKEND	DAYOWEEK: WEEKDAY	ROW TOTAL
-4 TO -1		58	144	202
		6	39.0	35.8
1 TO 3		101	153	254
		5	41.5	45.0
4 TO 11		37	72	109
		18.9	19.5	19.3
	COLUMN TOTAL	196	369	565
		34.7	65.3	100.0

CHI SQUARE = 6.09811 WITH 2 DEGREES OF FREEDOM
SIGNIFICANCE = 0.0474

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY FLORCOND FLOOR CONDITIONS

SUBPOS	COUNT / COL PCT	FLORCOND QUIET	FLORCOND NOISY	ROW TOTAL
-4 TO -1		139	63	202
		2	39.6	35.8
1 TO 3		180	74	254
		3	46.5	45.0
4 TO 11		87	22	109
		21.4	13.8	19.3
	COLUMN TOTAL	406	159	565
		71.9	28.1	100.0

CHI SQUARE = 4.46444 WITH 2 DEGREES OF FREEDOM
SIGNIFICANCE = 0.1073

NUMBER OF MISSING OBSERVATIONS = 35

C R O S S T A B U L A T I O N O F
SUBPOS- WRITTEN COMMENT SCORE
BY WEATHER

SUBPOS	COUNT / COL PCT	CLOUDY	RAIN	SUN	ROW TOTAL
-4 TO -1		74	20	108	202
		5	33.9	34.4	35.8
1 TO 3		79	24	151	254
		1	40.7	48.1	45.0
4 TO 11		39	15	55	109
		20.3	25.4	17.5	19.3
	COLUMN TOTAL	192	59	314	565
		34.0	10.4	55.6	100.0

CHI SQUARE = 3.94916 WITH 4 DEGREES OF FREEDOM
SIGNIFICANCE = 0.4129

NUMBER OF MISSING OBSERVATIONS = 35